

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

1902.657 Million cell updates/sec

Sequence: 1 DYKDDDDKALDTNYCFSSTE.....GRKPKVEQLSNMIVRSCKCS 120

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Maximum DB seq length: 2000000000

Listing first 45 summaries

```

1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	640	93.4	115	22	AAB73204		TGF-beta1 C-termin
2	640	93.4	115	24	ABG76033		Human TGF-beta 1 C
3	640	93.4	390	7	AAP61468		PreTGF-beta gene p
4	640	93.4	390	11	AAR04034		Sequence of pre-TG
5	640	93.4	390	11	AAR05258		Human pre-transfor
6	640	93.4	390	12	AAR13813		Human pro-TGF-beta
7	640	93.4	390	13	AAR20124		Sequence of simian
8	640	93.4	390	15	AAR46227		Human pre-TGF-beta
9	640	93.4	390	16	AAR73596		Human TGF-beta 1 p
10	640	93.4	390	17	AAR90827		Pre-transforming g
11	640	93.4	390	19	AAW78785		Human pre-transfor
12	640	93.4	390	22	AAM39186		Human polypeptide
13	640	93.4	390	23	AAU77101		Human transforming
14	640	93.4	390	23	AAE13596		Porcine transformi
15	640	93.4	390	23	AAE16943		Human transforming
16	640	93.4	390	24	ABB82781		TGFB1 Arg25Pro pol
17	640	93.4	391	9	AAP81362		Human transforming
18	640	93.4	391	16	AAR83054		Transforming growt
19	640	93.4	391	24	ABB82780		TGFB1 Arg25Pro pol
20	640	93.4	434	11	AAR03743		Monkey transformin
21	638	93.1	112	11	AAR04076		Sequence of mature
22	638	93.1	112	11	AAR08142		Platelet-derived h
23	638	93.1	112	14	AAR43263		TGF-beta. Homo sa
24	638	93.1	112	14	AAR42311		Recombinant human
25	638	93.1	112	17	AAR92773		Human TGF-beta 1.
26	638	93.1	112	17	AAR91956		Human transforming
27	638	93.1	112	18	AAW08173		TGF-beta1 active f
28	638	93.1	112	19	AAW78781		Human transforming
29	638	93.1	112	20	AAW08299		Human growth facto
30	638	93.1	112	20	AAW97091		The mature form of
31	638	93.1	112	20	AAW84207		Transforming growt
32	638	93.1	112	21	AAW92010		Human transforming
33	638	93.1	112	21	AAW67950		Human transforming
34	638	93.1	112	22	AAB35937		TGF-beta 1 amino a
35	638	93.1	112	23	AAM51939		Human TGFbeta prot
36	638	93.1	112	24	ABU08656		Human transforming
37	638	93.1	118	18	AAW08176		TGF-beta fusion pr
38	638	93.1	122	18	AAW08181		TGF-beta fusion pr
39	638	93.1	124	18	AAW08177		TGF-beta fusion pr
40	638	93.1	128	18	AAW08188		TGF-beta fusion pr
41	638	93.1	130	18	AAW08185		TGF-beta fusion pr
42	638	93.1	131	18	AAW08182		TGF-beta fusion pr
43	638	93.1	132	18	AAW08189		TGF-beta fusion pr
44	638	93.1	133	18	AAW08178		TGF-beta fusion pr
45	638	93.1	134	18	AAW08187		TGF-beta fusion pr

# ALIGNMENTS

RESULT 1

AAB73204

ID AAB73204 standard; Protein; 115 AA.

XX

AC AAB73204;

XX

DT 11-MAY-2001 (first entry)

XX

DE TGF-beta1 C-terminal sequence.

XX

KW Gene therapy; growth differentiation factor-8; GDF-8; AIDS; cachexia;

KW neurodegenerative disease; amyotrophic lateral sclerosis; obesity;

KW muscular dystrophy; musculodegenerative disease; tissue repair;

KW muscle wasting disease; neuromuscular disorder; spinal cord injury;

KW traumatic injury; congestive obstructive pulmonary disease.

XX

OS Unidentified.

XX

PN WO200112777-A2.

XX

PD 22-FEB-2001.

XX

PF 17-AUG-2000; 2000WO-US22884.

XX

PR 19-AUG-1999; 99US-0378238.

XX

PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

XX

PI Lee S, McPherron AC;

XX

DR WPI; 2001-211209/21.

XX

PT New substantially purified growth differentiation factor-8 polypeptide,

PT useful for treating muscle wasting disease, obesity, muscular

PT dystrophy, neuromuscular disorder, acquired immunodeficiency syndrome

PT and cachexia -

XX

PS Example 2; Fig 3; 124pp; English.

XX

CC The present invention relates to growth differentiation factor-8 (GDF-8)

CC coding sequences and proteins. The present sequence is a protein, which

CC was used in a sequence homology comparison with the GDF-8 protein

CC isolated in the present invention. GDF-8 is useful for treating

CC neurodegenerative diseases (e.g. amyotrophic lateral sclerosis and

CC muscular dystrophy), musculodegenerative diseases or in tissue repair due

CC to trauma, obesity and disorders related to abnormal proliferation of

CC adipocytes. GDF-8 is also useful for treating malignancies of the various

CC organ systems, particularly cells in muscle or adipose tissues and in

CC gene therapy for the treatment of cell proliferative or immunological

CC diseases mediated by GDF-8. In addition, GDF-8 is also useful for

CC treating muscle wasting disease, neuromuscular disorder, spinal cord

CC injury, traumatic injury, congestive obstructive pulmonary disease

CC (COPD), AIDS or cachexia.

XX

SQ Sequence 115 AA;

Query Match

93.4%; Score 640; DB 22; Length 115;

Best Local Similarity 99.1%; Pred. No. 9.5e-63;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      : ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 62

Qy     68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 115
```

RESULT 2

ABG76033

ID ABG76033 standard; Protein; 115 AA.

XX

AC ABG76033;

XX

DT 30-APR-2003 (first entry)

XX

DE Human TGF-beta 1 C-terminus.

XX

KW GDF-5; growth differentiation factor 5; TGF-beta; human;  
KW transforming growth factor beta; skeletal development; endometriosis;  
KW cartilage differentiation; cell proliferative disease; uterine tumour;  
KW bone dysplasia; spondylophyseal dysplasia; achondroplasia;  
KW dysplasia epiphysialis; metaphyseal dysostosis; hyperchondroplasia;  
KW enchondromatosis; hypophosphatasia; osteopetrosis; hyperphosphatasia;  
KW craniometaphyseal dysplasia; osteogenesis imperfecta; transgenic;  
KW idiopathic osteoporosis; Engelman's disease; TGF beta-1.

XX

OS Homo sapiens.

XX

PN US2002165361-A1.

XX

PD 07-NOV-2002.

XX

PF 12-JUN-2001; 2001US-0880708.

XX

PR 31-MAY-1995; 95US-0455559.

PR 01-SEP-1998; 98US-0145060.

PR 12-JAN-1993; 93US-0003144.

PR 12-JAN-1994; 94WO-US00657.

XX

PA (LEES/) LEE S.

PA (HUYN/) HUYNH T.

XX

PI Lee S, Huynh T;

XX

DR WPI; 2003-246676/25.

XX

PT New antibody specifically binding to a GDF-5 polypeptide, useful for  
PT diagnosing and treating cell proliferative disorders with aberrant  
PT GDF-5 activity, such as endometriosis, uterine tumors and those  
PT involving skeletal tissues -

XX

PS Example 2; Fig 3A; 36pp; English.



XX  
CC The invention relates to an antibody that specifically binds to growth  
CC differentiation factor-5 (GDF-5, a member of the TGF-Beta (transforming  
CC growth factor beta) superfamily of proteins) polypeptide appearing  
CC as ABG76018. In order to determine the biological activity of GDF-5 in  
CC vivo, transgenic mice were constructed that express GDF-5 ectopically.  
CC Analysis of two independent transgenic mouse lines showed that these  
CC animals have ectopic bone formation with evident muscle tissue. This  
CC showed that GDF-5 was capable of inducing bone formation in vivo.  
CC The antibody is useful for the diagnosis and treatment of cell  
CC proliferative disorders associated with aberrant GDF-5 activity, such  
CC as endometriosis, uterine tumors, those involving skeletal tissues,  
CC endometriosis, cartilage differentiation, cell proliferative disease,  
CC uterine tumour, bone dysplasia, spondylopathy, dysplasia,  
CC achondroplasia, dysplasia epiphysealis, metaphyseal dysostosis,  
CC hyperchondroplasia, enchondromatosis, hypophosphatasia, osteopetrosis,  
CC hyperphosphatasia, craniometaphyseal dysplasia, osteogenesis imperfecta,  
CC idiopathic osteoporosis and Engelmann's disease. The present sequence  
CC represents a member of the TGFbeta superfamily used to determine  
CC regions of sequence similarity for design of degenerate primers for  
CC isolation of GDF-5.

XX  
SQ Sequence 115 AA;

Query Match 93.4%; Score 640; DB 24; Length 115;  
Best Local Similarity 99.1%; Pred. No. 9.5e-63;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 3 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 62

Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 63 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 115

### RESULT 3

AAP61468

ID AAP61468 standard; Protein; 390 AA.

XX

AC AAP61468;

XX

DT 31-OCT-2002 (updated)

DT 28-OCT-1991 (first entry)

XX

DE PreTGF-beta gene product.

XX

KW Transforming growth factor beta; cancer; wound healing.

XX

OS Unidentified.

XX

FH Key Location/Qualifiers

FT Protein 279..390

XX

PN EP200341-A.

XX

PD 10-DEC-1986.  
 XX  
 PF 21-MAR-1986; 86EP-0302112.  
 XX  
 PR 22-MAR-1985; 85US-0715142.  
 PR 13-MAR-1987; 87US-0025423.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Derynck RMA;  
 XX  
 DR WPI; 1986-326875/50.  
 DR N-PSDB; AAN60972.  
 XX  
 PT TGF-beta prodn. from transformed hosts - useful esp. for treating  
 PT wounds (J6 2/9/86).  
 XX  
 PS Disclosure; Fig 1b; 26pp; English.  
 XX  
 CC The gene product is known to stimulate cell proliferation and  
 CC inhibit anchorage-dependent growth of a variety of human cancer cell  
 CC lines, it is esp. useful in treatment of burns and the promotion of  
 CC surface and internal wound healing. TGF-beta may be expressed from a  
 CC transformed CHO cell line.  
 CC (Updated on 31-OCT-2002 to add missing OS field.)  
 XX  
 SQ Sequence 390 AA;  
  
 Query Match 93.4%; Score 640; DB 7; Length 390;  
 Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
 Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

#### RESULT 4

AAR04034

ID AAR04034 standard; protein; 390 AA.  
 XX  
 AC AAR04034;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 31-OCT-2002 (updated)  
 DT 31-MAY-1989 (first entry)  
 XX  
 DE Sequence of pre-TGF-beta 1.  
 XX  
 KW Transforming growth factor beta-3 (TGF beta 3); tumour cells; growth  
 KW inhibition.  
 XX  
 OS Homo sapiens.

XX  
 FH Key Location/Qualifiers  
 FT CDS 348..500  
 XX  
 PN WO8912101-A.  
 XX  
 PD 14-DEC-1989.  
 XX  
 PF 08-JUN-1988; 88WO-US01945.  
 XX  
 PR 08-JUN-1988; 88WO-US01945.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Dernyck RMA, Goeddel DV;  
 XX  
 DR WPI; 1990-007474/01.  
 DR N-PSDB; AAQ02815.  
 XX  
 PT Nucleotide sequence encoding transforming growth factor beta-3 used as a  
 PT probe, or to produce TGF beta 3, for inhibiting growth of certain normal  
 PT and neoplastic cells, eg A549.  
 XX  
 PS Disclosure; Fig. 2; 61pp; English.  
 XX  
 CC Sequence is an exon of transforming growth factor-beta 1 (pre-TGF-beta  
 CC 1) polypeptide and corresponds to AA's 288-338 of mature TGF-beta 1. The  
 CC nucleic acid encoding second subtype of TGF-beta (TGF-beta 3) is useful  
 CC as a probe or to produce TGF-beta 3 for inhibition of normal and  
 CC neoplastic cell growth.  
 CC (Updated on 31-OCT-2002 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 11; Length 390;  
 Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
 Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : |||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
 QY 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
 |||||  
 Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

# RESULT 5

AAR05258

ID AAR05258 standard; protein; 390 AA.

XX

AC AAR05258;

XX

DT 25-MAR-2003 (updated)

DT 05-AUG-1990 (first entry)

XX  
 DE Human pre-transforming growth factor-beta-1 (pre-TGF-beta-1).  
 XX  
 KW Transforming growth factor-beta-1 (TGF-beta-1);  
 KW neoplastic cell line inhibition;  
 KW EGF-potentiated anchorage-independent growth.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..278  
 FT Protein 279..2011  
 FT Domain 8..23  
 FT /note="hydrophobic domain"  
 FT Modified-site 82..84  
 FT /note="potential N-glycosylation site "  
 FT Modified-site 136..138  
 FT /note="as above"  
 FT Modified-site 176..178  
 FT /note="as above"  
 FT Cleavage-site 277..278  
 FT /note="proteolytic cleavage site"  
 XX  
 PN US4886747-A.  
 XX  
 PD 12-DEC-1989.  
 XX  
 PF 13-MAR-1987; 87US-0025423.  
 XX  
 PR 13-MAR-1987; 87US-0025423.  
 PR 22-MAR-1985; 85US-0715142.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Derynck RMA, Goeddel DV;  
 XX  
 DR WPI; 1990-051338/07.  
 DR N-PSDB; AAQ93301.  
 XX  
 PT Nucleic acid encoding transforming growth factor-beta -  
 PT cloned into expression vectors for expression in eukaryotic host  
 PT cells for therapeutic use  
 XX  
 PS Disclosure; Fig 1b; 28pp; English.  
 XX  
 CC Pre-TGF-beta-1 is a polypeptide of 390 amino acids. Post-translational  
 CC cleavage of the precursor gives rise to the mature TGF-beta monomer.  
 CC The sequence for human TGF-beta was determined by direct amino acid  
 CC sequence analysis and by deduction from the TGF-beta cDNA. It is  
 CC capable of inducing EGF-potentiated anchorage-independent growth of  
 CC target cell lines, and/or growth inhibition of neoplastic cell lines. It  
 CC can be used for treating wounds, eg burns or epidermal ulcers.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 390 AA;

Query Match

93.4%; Score 640; DB 11; Length 390;

Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy     68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    338 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 6

AAR13813

ID AAR13813 standard; Protein; 390 AA.

XX

AC AAR13813;

XX

DT 20-NOV-1991 (first entry)

XX

DE Human pro-TGF-beta 1.

XX

KW Osteogenetic; tumoricidal.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..29

FT /note= "signal peptide"

FT Peptide 30..390

FT /note= "pro-TGF-beta 1"

FT Peptide 279..390

FT /note= "TGF-beta 1"

XX

PN JP03180192-A.

XX

PD 06-AUG-1991.

XX

PF 07-DEC-1989; 89JP-0318243.

XX

PR 07-DEC-1989; 89JP-0318243.

XX

PA (KIRI ) KIRIN BREWERY KK.

XX

DR WPI; 1991-271579/37.

DR N-PSDB; AAQ13392.

XX

PT Human pro-TGF-beta 1 prodn., for osteo-genetic activity - by

PT preparing DNA chain contg. base sequence coding for human

PT pre:pro-TGF-beta 1, forming expression vector etc.

XX

PS Claim 1; Fig 1; 16pp; Japanese.

XX

CC The amino acid sequence codes for human prepro-TGF-beta 1 which

CC can be produced by recombinant methods, it has osteogenetic and

CC tumoricidal activity.

XX

SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 12; Length 390;  
Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      :|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

QY     68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    338 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 7

AAR20124

ID AAR20124 standard; Protein; 390 AA.

XX

AC AAR20124;

XX

DT 25-MAR-2003 (updated)

DT 16-APR-1992 (first entry)

XX

DE Sequence of simian transforming growth factor (TGF) beta-1.

XX

KW Hypertension therapy; hypotensive agent; blood pressure modulator.

XX

OS Monkey.

XX

FH Key Location/Qualifiers

FT Peptide 8..21

FT Protein 279..390

XX

PN WO9119513-A.

XX

PD 26-DEC-1991.

XX

PF 20-JUN-1991; 91WO-US04449.

XX

PR 20-JUN-1990; 90US-0541221.

XX

PA (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX

PI Oleson FB, Comerreski CR;

XX

DR WPI; 1992-024199/03.

DR N-PSDB; AAQ20289.

XX

PT Use of transforming growth factor (TGF)-beta and their  
PT antagonists - for modulating blood pressure, for treating  
PT hypertension and hypotension

XX

PS Disclosure; Fig 1; 42pp; English.

XX

CC A new method for treating hypertension comprises administering a  
CC transforming growth factor (TGF)-beta to an individual at a dose

CC effective for lowering blood pressure; the TGF-beta may be e.g.  
 CC mature TGF-beta, TGF-beta2, a mature TGF-beta1/beta2 hybrid, TGF-  
 CC beta1 precursor, a latent TGF-beta2 precursor, hybrid TGF-beta1/TGF-  
 CC beta2 precursor, a latent TGF-beta1 complex or a latent TGF-beta2  
 CC complex.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 13; Length 390;  
 Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
 Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 8

AAR46227

ID AAR46227 standard; Protein; 390 AA.

XX

AC AAR46227;

XX

DT 25-MAR-2003 (updated)

DT 09-JUL-1994 (first entry)

XX

DE Human pre-TGF-beta-1.

XX

KW TGF-beta-1; TGF-beta-2; transforming growth factor beta-1;

KW transforming growth factor beta-3; recombinant; wound healing;

KW vulnerary.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 279..390

FT /label= Mat\_peptide

FT Cleavage-site 279

FT /note= "TGF-beta-1 release site"

FT Modified-site 82..84

FT /label= N-glycosylation\_site

FT Modified-site 136..138

FT /label= N-glycosylation\_site

FT Modified-site 176..178

FT /label= N-glycosylation\_site

XX

PN US5284763-A.

XX

PD 08-FEB-1994.

XX

PF 04-MAR-1992; 92US-0845893.

XX

PR 22-MAR-1985; 85US-0715142.  
PR 13-MAR-1987; 87US-0025423.  
PR 04-AUG-1989; 89US-0389929.  
PR 04-MAR-1992; 92US-0845893.

XX

PA (GETH ) GENENTECH INC.

XX

PI Derynk RMA, Goeddel DV;

XX

DR WPI; 1994-056343/07.

DR N-PSDB; AAQ56923.

XX

PT Nucleic acid sequences encoding transforming growth factor-beta -  
PT diagnostic probes, and for use in therapeutics

XX

PS Disclosure; Fig 1b; 25pp; English.

XX

CC cDNA sequences were determined for human pre-TGF-beta-1 (AAQ56923),  
CC pig TGF-beta-3 (AAQ56925) and human TGF-beta-3 (AAQ56926), and the  
CC corresponding amino acid sequences were determined (AAR46227-29,  
CC respectively). A genomic fragment corresponding to a human TGF-  
CC beta-1 exon (AAQ56924) was also isolated and its amino acid sequence  
CC determined (AAR46230). The sequences have been used in the  
CC construction of vectors for the expression of recombinant TGF-  
CC beta.

CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 15; Length 390;  
Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
|||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

# RESULT 9

AAR73596

ID AAR73596 standard; Protein; 390 AA.

XX

AC AAR73596;

XX

DT 25-MAR-2003 (updated)

DT 20-DEC-1995 (first entry)

XX

DE Human TGF-beta 1 protein.

XX

KW Transforming growth factor-beta; Human TGF-beta protein; TGF-beta 1;  
KW TGF-beta 3; osteogenic cell source; OCS; bone deficiency;  
KW bone-inducing cofactor.

XX



OS Homo sapiens.  
 XX  
 PN US5409896-A.  
 XX  
 PD 25-APR-1995.  
 XX  
 PF 12-NOV-1993; 93US-0132405.  
 XX  
 PR 01-SEP-1989; 89US-0401906.  
 PR 12-NOV-1991; 91US-0790856.  
 PR 18-MAY-1993; 93US-0063841.  
 PR 12-NOV-1993; 93US-0132405.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Ammann AJ, Rudman CG;  
 XX  
 DR WPI; 1995-169610/22.  
 XX  
 PT Compsn. for treating skeletal tissue deficiency - comprising  
 PT transforming growth factor-beta and an osteogenic cell source in a  
 PT carrier  
 XX  
 PS Claim 3; Column 15-18; 19pp; English.  
 XX  
 CC This sequence represents human transforming growth factor-beta 1  
 CC (TGF-beta 1). This sequence and the sequence for human TGF-beta 3 (see  
 CC AAR73598) are claimed within the scope of the invention. The invention  
 CC is a composition consisting of a TGF-beta protein and an osteogenic cell  
 CC source (OCS) formulated in an acceptable carrier other than a bone  
 CC morphogenic cofactor. This composition can be used for the restoration  
 CC of bone deficiency. This provides for the generation of mature bone  
 CC only where it is required, without the inclusion of a specific  
 CC bone-inducing cofactor. This method can be used with any of the 5 human  
 CC TGF-beta's or with TGF-beta from other species.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 16; Length 390;  
 Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
 Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 10  
 AAR90827  
 ID AAR90827 standard; Protein; 390 AA.  
 XX  
 AC AAR90827;

XX  
 DT 25-MAR-2003 (updated)  
 DT 25-JAN-1980 (first entry)  
 XX  
 DE Pre-transforming growth factor beta 1.  
 XX  
 KW transforming growth factor beta 1; wound healing;  
 KW recombinant production.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 8..23  
 FT /note= "hydrophobic domain"  
 FT Modified-site 82..84  
 FT /note= "potential N-glycosylation site"  
 FT Modified-site 136..138  
 FT /note= "potential N-glycosylation site"  
 FT Modified-site 176..178  
 FT /note= "potential N-glycosylation site"  
 FT Cleavage-site 277..279  
 FT /note= "trypsin-like peptidase cleavage site"  
 FT Protein 279..390  
 FT /label= mature\_TGF\_beta\_1  
 XX  
 PN US5482851-A.  
 XX  
 PD 09-JAN-1996.  
 XX  
 PF 05-NOV-1993; 93US-0147364.  
 XX  
 PR 13-MAR-1987; 87US-0025423.  
 PR 22-MAR-1985; 85US-0715142.  
 PR 04-AUG-1989; 89US-0389929.  
 PR 04-MAR-1992; 92US-0845893.  
 PR 05-NOV-1993; 93US-0147364.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Derynck RMA, Goeddel DV;  
 XX  
 DR WPI; 1996-076891/08.  
 DR N-PSDB; AAT15720.  
 XX  
 PT New recombinant human transforming growth factor-beta prods. - produced  
 PT using Chinese hamster ovary cells, for use in diagnostic applications  
 PT or in therapy  
 XX  
 PS Example 3; Fig 1A-C; 26pp; English.  
 XX  
 CC The pre-transforming growth factor (TGF) beta 1 protein is encoded  
 CC by AAT15720. The mature TGF beta 1 monomer is cleaved from the  
 CC precursor at the Arg-Arg dipeptide immediately preceding the mature  
 CC TGF-beta 1 NH2-terminus. It does not contain a recognisable N-terminal  
 CC signal peptide typical of most secreted proteins. The pre-TGF beta 1  
 CC contains several pairs of basic residues which could undergo  
 CC post-translational cleavage and give rise to separate polypeptide

CC entities. The precursor contains 3 potential N-glycosylation sites, none  
 CC of which are localised in the mature TGF beta 1. This is useful in  
 CC purification of the mature protein. TGF beta 1 can be used in, e.g. wound  
 CC healing.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 17; Length 390;  
 Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
 Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 11

AAW78785

ID AAW78785 standard; Protein; 390 AA.

XX

AC AAW78785;

XX

DT 25-MAR-2003 (updated)

DT 21-DEC-1998 (first entry)

XX

DE Human pre-transforming growth factor-beta 1.

XX

KW Transforming growth factor-beta 1; TGF-beta 1; human.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 8..23

FT /note= "hydrophobic domain"

FT Protein 279..390

FT /label= Mat\_protein

FT Modified-site 82..84

FT /note= "Asn is N-glycosylated"

FT Modified-site 136..138

FT /note= "Asn is N-glycosylated"

FT Modified-site 176..178

FT /note= "Asn is N-glycosylated"

FT Cleavage-site 277..278

FT /note= "cleavage site for release of TGF-beta 1"

XX

PN US5801231-A.

XX

PD 01-SEP-1998.

XX

PF 30-MAY-1995; 95US-0454468.

XX

PR 13-MAR-1987; 87US-0025423.

PR 22-MAR-1985; 85US-0715142.  
PR 04-AUG-1989; 89US-0389929.  
PR 04-MAR-1992; 92US-0845893.  
PR 05-NOV-1993; 93US-0147364.  
PR 30-MAY-1995; 95US-0454468.

XX

PA (GETH ) GENENTECH INC.

XX

PI Derynck RMA, Goeddel DV;

XX

DR WPI; 1998-494840/42.

DR N-PSDB; AAV52933.

XX

PT DNA encoding transforming growth factor-beta precursor sequence -  
PT useful for analysis to perform manipulations to increase yield of  
PT recombinant production of the protein

XX

PS Example 3; Fig 1B 1-3; 26pp; English.

XX

CC This is the amino acid sequence of human transforming growth  
CC factor-beta 1 precursor (preTGF-beta 1). It was deduced from  
CC a preTGF-beta 1 cDNA sequence (see AAV52933). The invention relates  
CC to the recombinant production of TGF-beta. Biologically active  
CC TGF-beta is defined as being capable of inducing EGF-potentiated  
CC anchorage independent growth of target cell lines and/or growth  
CC inhibition of neoplastic cell lines. Nucleic acids encoding  
CC TGF-beta have been isolated and cloned into vectors which are  
CC replicated in bacteria and expressed in eukaryotic cells. TGF-beta  
CC recovered from transformed cells is used in known therapeutic  
CC applications.

CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 19; Length 390;  
Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
:||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

## RESULT 12

AAM39186

ID AAM39186 standard; Protein; 390 AA.

XX

AC AAM39186;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 2331.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR N-PSDB; AAI58342.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Example 4; SEQ ID NO 2331; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 22; Length 390;  
Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy     68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 13

AAU77101

ID AAU77101 standard; Protein; 390 AA.

XX

AC AAU77101;

XX

DT 05-JUN-2002 (first entry)

XX

DE Human transforming growth factor beta 1 (TGF-beta-1) polypeptide.

XX

KW Human; transforming growth factor beta; TGF-beta; insulin production;  
KW type I diabetes mellitus; pancreatic cell outgrowth; wound healing;  
KW pancreatic duct tissue; ischaemia; stroke; nervous system aging;  
KW neurological condition; neurodegenerative disease; inflammation;  
KW vascular injury; chemical injury; traumatic injury; tumour-induced injury;  
KW amyotrophic lateral sclerosis; spinocerebellar degeneration;  
KW immunological disease; multiple sclerosis; TGF-beta-1.

XX

OS Homo sapiens.

XX

PN WO200212336-A2.

XX

PD 14-FEB-2002.

XX

PF 09-FEB-2001; 2001WO-US04192.

XX

PR 09-AUG-2000; 2000US-0635368.

XX

PA (CURI-) CURIS INC.

XX

PI Wang M, Pang K;

XX

DR WPI; 2002-257468/30.

XX

PT Treating a subject with a disorder resulting from insufficient insulin  
PT production, and inducing outgrowth of pancreatic cells, involves using  
PT a transforming growth factor beta therapeutic -

XX

PS Disclosure; Fig 1; 77pp; English.

XX

CC The invention relates to treating a subject with a disorder resulting  
CC from insufficient insulin production, involving contacting the subject  
CC with a transforming growth factor beta (TGF-beta) therapeutic. TGF-beta  
CC polypeptides can be used for treating a subject with a disorder resulting

CC from insufficient insulin production, e.g. type I diabetes mellitus, and  
 CC for inducing outgrowth of pancreatic cells associated with pancreatic  
 CC duct tissue within a subject. A composition comprising a TGF-beta protein  
 CC may be useful in wound healing and treatment of neurological conditions  
 CC derived from acute, subacute or chronic injury to the nervous system,  
 CC including traumatic injury, chemical injury, vasal injury and deficits  
 CC (such as ischaemia resulting from stroke), together with  
 CC infectious/inflammatory and tumour-induced injury, aging of the nervous  
 CC system including Alzheimer's disease, chronic neurodegenerative diseases  
 CC including Parkinson's disease, Huntington's chorea, amyotrophic lateral  
 CC sclerosis, spinocerebellar degenerations and chronic immunological  
 CC diseases of the nervous system or affecting the nervous system, including  
 CC multiple sclerosis. This sequence represents the human TGF-beta-1  
 CC protein.  
 XX  
 SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 23; Length 390;  
 Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
 Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : |||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 |||||  
 Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

#### RESULT 14

AAE13596

ID AAE13596 standard; Protein; 390 AA.

XX

AC AAE13596;

XX

DT 26-FEB-2002 (first entry)

XX

DE Porcine transforming growth factor beta 1 (TGF-beta1) mutant.

XX

KW Porcine; transforming growth factor beta 1; TGF-beta1; gene therapy;  
 KW IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;  
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
 KW diabetes mellitus; sarcoidosis; psoriasis; dermatological; mutant;  
 KW mutein.

XX

OS Sus scrofa.

XX

FH Key Location/Qualifiers

FT Misc-difference 223

FT /note= "Wild type Cys substituted with Ser"

FT Misc-difference 225

FT /note= "Wild type Cys substituted with Ser"

XX

PN WO200181404-A2.

XX

PD 01-NOV-2001.

XX  
PF 20-APR-2001; 2001WO-US12980.  
XX  
PR 20-APR-2000; 2000US-199014P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Strober W, Nakamura K, Kitani A, Fuss IJ;  
XX  
DR WPI; 2002-026155/03.  
DR N-PSDB; AAD22696.  
XX  
PT Composition for treating autoimmune diseases e.g. inflammatory bowel  
PT disease in humans, comprises vector containing transforming growth  
PT factor-beta under the control of inducible promoter -  
XX  
PS Example 1; Fig 1; 78pp; English.  
XX  
CC The invention relates to a composition containing a vector comprising a  
CC gene encoding a regulatory transcription factor under the control of a  
CC promoter encoding a transforming growth factor-beta (TGF-beta). The  
CC vector is useful for expressing TGF-beta, such as TGF-beta1, TGF-beta2  
CC or TGF-beta3, its variants or homologues, by transfecting a cell which  
CC is part of a host suspected of having an autoimmune disease, especially  
CC inflammatory bowel disease (IBD), under conditions such that the  
CC polypeptide encoded by the nucleic acid sequence in the vector is  
CC expressed. The vector is delivered using a delivery system. The delivery  
CC of the vector results in substantial elimination of symptoms of the  
CC autoimmune disease and increased production of IL-10 by the host. The  
CC composition is useful for treating various diseases with an autoimmune  
CC component such as multiple sclerosis, rheumatoid arthritis, systemic  
CC lupus erythematosus, insulin-dependent diabetes mellitus, sarcoidosis  
CC and psoriasis, and also for assaying the expression of a gene in a cell.  
CC The vector is further useful for screening of the effect of test  
CC compounds on cytokine (e.g. TGF-beta) expression of transfected cells.  
CC The present sequence is porcine TGF-beta1 mutant.  
XX  
SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 23; Length 390;  
Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
: ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
Qy 68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 338 KVLALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 15  
AAE16943  
ID AAE16943 standard; Protein; 390 AA.  
XX  
AC AAE16943;



XX  
DT 18-APR-2002 (first entry)  
XX  
DE Human transforming growth factor-beta1 (TGF-beta1) protein.  
XX  
KW Human; transforming growth factor-beta1; TGF-beta1; osteoporosis;  
KW latency associated peptide; LAP; integrin alphavbeta3; apoptosis;  
KW immunomodulation; inflammatory disease; fibrotic disease; cancer;  
KW diabetic retinopathy; chronic obstructive pulmonary disorder;  
KW bone resorption; rheumatoid arthritis; psoriasis; restenosis;  
KW atherosclerosis; liver fibrosis; asthma; cytostatic; osteopathic;  
KW ophthalmological; antiarteriosclerotic; vasotropic.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..29  
FT /label= Signal\_peptide  
FT Region 30..278  
FT /note= "LAP-beta1"  
FT Domain 244..246  
FT /note= "RGD motif"  
FT Protein 279..390  
FT /note= "Human mature TGF-beta1 protein"  
XX  
PN WO200190186-A1.  
XX  
PD 29-NOV-2001.  
XX  
PF 25-MAY-2001; 2001WO-GB02352.  
XX  
PR 26-MAY-2000; 2000GB-0012991.  
PR 05-JAN-2001; 2001GB-0000286.  
XX  
PA (GLAX ) GLAXO GROUP LTD.  
XX  
PI Ludbrook S, Barry S, Horgan C, Miller D;  
XX  
DR WPI; 2002-097645/13.  
XX  
PT Identifying modulators of interactions between latency associated  
PT peptides and integrin alphavbeta3 for therapeutics, by contacting the  
PT peptide and integrin with a test product and determining if the product  
PT modulates interaction -  
XX  
PS Disclosure; Page 37-39; 44pp; English.  
XX  
CC The invention relates to a method for identification of a modulator  
CC of the interaction between latency associated peptide (LAP) of  
CC transforming growth factor-beta1 (TGF-beta1) and integrin alphavbeta3.  
CC The method is useful for identifying a modulator of the interaction  
CC between LAP and integrin alphavbeta3. The method is useful for  
CC immunomodulation, in the treatment of inflammatory disease, fibrotic  
CC disease, cancer, diabetic retinopathy, bone resorption or osteoporosis,  
CC and for preventing apoptosis administering the modulator to the host.  
CC The modulator (inhibitor of the interaction between LAP-beta1 and  
CC integrin alphavbeta3) is useful in the manufacture of a medicament for

CC immunomodulation. The modulator (activator of the interaction between  
CC LAP-betal and integrin alphavbeta3 ) is useful in the manufacture of  
CC medicament for preventing apoptosis. The modulator is useful for  
CC treating a inflammatory or fibrotic disease such as chronic obstructive  
CC pulmonary disorder, rheumatoid arthritis, psoriasis, restenosis,  
CC atherosclerosis, liver fibrosis and asthma. The present sequence is  
CC human TGF-betal protein.

XX

SQ Sequence 390 AA;

Query Match 93.4%; Score 640; DB 23; Length 390;  
Best Local Similarity 99.1%; Pred. No. 3.8e-62;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

Search completed: October 28, 2003, 09:06:46  
Job time : 11.0108 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 09:09:54 ; Search time 7.28061 Seconds  
(without alignments)  
2760.110 Million cell updates/sec

Title: US-10-017-372E-11  
Perfect score: 685  
Sequence: 1 DYKDDDDKALDTNYCFSSTE.....GRKPKVEQLSNMIVRSCKCS 120

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*

```

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	640	93.4	114	10	US-09-813-459-22	Sequence 22, Appl
2	640	93.4	114	14	US-10-115-406-21	Sequence 21, Appl
3	640	93.4	114	15	US-10-154-333-23	Sequence 23, Appl
4	640	93.4	115	10	US-09-859-211-47	Sequence 47, Appl
5	640	93.4	115	10	US-09-880-708-25	Sequence 25, Appl
6	640	93.4	115	11	US-09-872-856-47	Sequence 47, Appl
7	640	93.4	115	15	US-10-335-483-29	Sequence 29, Appl
8	640	93.4	139	14	US-10-002-278-8	Sequence 8, Appli
9	640	93.4	315	11	US-09-214-592-25	Sequence 25, Appl
10	640	93.4	390	10	US-09-756-283A-23	Sequence 23, Appl
11	640	93.4	390	11	US-09-214-592-26	Sequence 26, Appl
12	640	93.4	390	11	US-09-214-592-28	Sequence 28, Appl
13	640	93.4	390	11	US-09-214-592-29	Sequence 29, Appl
14	640	93.4	390	11	US-09-214-592-33	Sequence 33, Appl
15	640	93.4	390	12	US-10-276-947-1	Sequence 1, Appli
16	640	93.4	390	15	US-10-087-268-2	Sequence 2, Appli
17	640	93.4	390	15	US-10-087-268-5	Sequence 5, Appli
18	640	93.4	391	11	US-09-214-592-17	Sequence 17, Appl
19	638	93.1	112	10	US-09-813-271B-2	Sequence 2, Appli
20	638	93.1	113	10	US-09-813-398-13	Sequence 13, Appl
21	637	93.0	390	11	US-09-214-592-20	Sequence 20, Appl
22	637	93.0	390	11	US-09-214-592-23	Sequence 23, Appl
23	562	82.0	98	12	US-10-187-394-1	Sequence 1, Appli
24	542	79.1	116	14	US-10-115-406-24	Sequence 24, Appl
25	542	79.1	116	15	US-10-154-333-26	Sequence 26, Appl
26	542	79.1	373	11	US-09-214-592-32	Sequence 32, Appl
27	535	78.1	112	10	US-09-813-271B-8	Sequence 8, Appli
28	531	77.5	304	10	US-09-756-283A-26	Sequence 26, Appl
29	505	73.7	114	14	US-10-115-406-25	Sequence 25, Appl
30	505	73.7	114	15	US-10-154-333-27	Sequence 27, Appl
31	505	73.7	382	11	US-09-214-592-34	Sequence 34, Appl

32	504	73.6	112	10	US-09-813-271B-12	Sequence 12, Appl
33	503	73.4	383	10	US-09-756-283A-27	Sequence 27, Appl
34	502	73.3	114	10	US-09-813-459-24	Sequence 24, Appl
35	502	73.3	114	14	US-10-115-406-23	Sequence 23, Appl
36	502	73.3	114	15	US-10-154-333-25	Sequence 25, Appl
37	502	73.3	115	10	US-09-880-708-27	Sequence 27, Appl
38	502	73.3	410	11	US-09-214-592-22	Sequence 22, Appl
39	502	73.3	412	11	US-09-214-592-19	Sequence 19, Appl
40	502	73.3	412	14	US-10-028-158-21	Sequence 21, Appl
41	501	73.1	412	11	US-09-214-592-24	Sequence 24, Appl
42	500	73.0	112	10	US-09-813-271B-6	Sequence 6, Appli
43	500	73.0	113	10	US-09-813-398-15	Sequence 15, Appl
44	499	72.8	98	12	US-10-187-394-4	Sequence 4, Appli
45	499	72.8	412	11	US-09-214-592-31	Sequence 31, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-813-459-22

; Sequence 22, Application US/09813459

; Patent No. US20020107369A1

; GENERAL INFORMATION:

; APPLICANT: Lee, Se-Jin

; Cunningham, No. US20020107369A1

; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-10

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Spensley Horn Jubas & Lubitz

; STREET: 1880 Century Park East, Suite 500

; CITY: Los Angeles

; STATE: California

; COUNTRY: USA

; ZIP: 90067

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/813,459

; FILING DATE: 20-Mar-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/624,635

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Wetherell, Jr., Ph.D., John R.,

; REGISTRATION NUMBER: 31,678

; REFERENCE/DOCKET NUMBER: PD-3054

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 455-5100

; TELEFAX: (619) 455-5110

; INFORMATION FOR SEQ ID NO: 22:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 114 amino acids

; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; IMMEDIATE SOURCE:  
; CLONE: TGF-beta-1  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..114  
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-813-459-22

Query Match 93.4%; Score 640; DB 10; Length 114;  
Best Local Similarity 99.1%; Pred. No. 8e-60;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 2 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 61  
  
Qy 68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 62 KVLALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 114

RESULT 2

US-10-115-406-21

; Sequence 21, Application US/10115406  
; Publication No. US20020127612A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: LEE, Se-Jin  
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-9  
; FILE REFERENCE: JHU1190-3  
; CURRENT APPLICATION NUMBER: US/10/115,406  
; CURRENT FILING DATE: 2002-04-02  
; PRIOR APPLICATION NUMBER: 09/301,520  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: US 09/172,062  
; PRIOR FILING DATE: 1998-10-13  
; PRIOR APPLICATION NUMBER: US 08/491,835  
; PRIOR FILING DATE: 1995-10-23  
; PRIOR APPLICATION NUMBER: PCT/US94/00685  
; PRIOR FILING DATE: 1994-01-12  
; PRIOR APPLICATION NUMBER: US 08/003,303  
; PRIOR FILING DATE: 1993-01-12  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 21  
; LENGTH: 114  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-115-406-21

Query Match 93.4%; Score 640; DB 14; Length 114;  
Best Local Similarity 99.1%; Pred. No. 8e-60;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 2 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 61

Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 62 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 114

RESULT 3

US-10-154-333-23

; Sequence 23, Application US/10154333

; Publication No. US20030109684A1

; GENERAL INFORMATION:

; APPLICANT: JOHNS HOPKINS UNIVERSITY

; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-3

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ

; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR

; CITY: LOS ANGELES

; STATE: CALIFORNIA

; COUNTRY: US

; ZIP: 90067

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/154,333

; FILING DATE: 21-May-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/389,705

; FILING DATE: 03-Sep-1999

; APPLICATION NUMBER: 09/153,733

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: WETHERELL, JR. Ph.D., JOHN R.

; REGISTRATION NUMBER: 31,678

; REFERENCE/DOCKET NUMBER: FD2279 PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 455-5100

; TELEFAX: (619) 455-5110

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 114 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; IMMEDIATE SOURCE:

; CLONE: TGF-beta 1

; FEATURE:

; NAME/KEY: Protein

; LOCATION: 1..114  
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-10-154-333-23

Query Match 93.4%; Score 640; DB 15; Length 114;  
Best Local Similarity 99.1%; Pred. No. 8e-60;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
:||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 2 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 61  
  
Qy 68 KVLALYNQHNPASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
:||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 62 KVLALYNQHNPASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 114

RESULT 4

US-09-859-211-47  
; Sequence 47, Application US/09859211  
; Patent No. US20020157125A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Se-Jin  
; APPLICANT: McPherron, Alexandra C.  
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-8  
; FILE REFERENCE: 07265/144001  
; CURRENT APPLICATION NUMBER: US/09/859,211  
; CURRENT FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: 09/019,070  
; PRIOR FILING DATE: 1998-02-05  
; PRIOR APPLICATION NUMBER: 08/862,445  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 08/847,910  
; PRIOR FILING DATE: 1997-04-28  
; PRIOR APPLICATION NUMBER: 08/795,071  
; PRIOR FILING DATE: 1997-02-05  
; PRIOR APPLICATION NUMBER: 08/525,596  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: PCT/US94/03019  
; PRIOR FILING DATE: 1994-03-18  
; PRIOR APPLICATION NUMBER: 08/033,923  
; PRIOR FILING DATE: 1993-03-19  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 47  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-859-211-47

Query Match 93.4%; Score 640; DB 10; Length 115;  
Best Local Similarity 99.1%; Pred. No. 8e-60;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
:||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 3 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 62

Qy 68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 63 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 115

RESULT 5

US-09-880-708-25

; Sequence 25, Application US/09880708

; Patent No. US20020165361A1

; GENERAL INFORMATION:

; APPLICANT: Lee, Se-Jin

; Huynh, Thanh

; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-5

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Gray Cary Ware & Freidenrich LLP

; STREET: 4365 Executive Drive, Suite 1600

; CITY: San Diego

; STATE: CA

; COUNTRY: USA

; ZIP: 92121-2189

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/880,708

; FILING DATE: 12-Jun-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/145,060

; FILING DATE: <Unknown>

; APPLICATION NUMBER: 08/003,144

; FILING DATE: 12-JAN-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Lisa A. Haile, Ph.D.

; REGISTRATION NUMBER: 38,347

; REFERENCE/DOCKET NUMBER: 07265/057002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 858/677-1456

; TELEFAX: 619/677-1465

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 115 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; IMMEDIATE SOURCE:

; CLONE: TGF-beta-1

; SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-09-880-708-25

Query Match 93.4%; Score 640; DB 10; Length 115;

Best Local Similarity 99.1%; Pred. No. 8e-60;

Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : |||||  
 Db 3 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 62

Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 |||||  
 Db 63 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 115

RESULT 6

US-09-872-856-47

; Sequence 47, Application US/09872856  
 ; Publication No. US20030074680A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Johns Hopkins University School of Medicine  
 ; APPLICANT: Lee, Se-Jin  
 ; APPLICANT: McPherron, Alexandra  
 ; TITLE OF INVENTION: Growth Differentiation Factor-8  
 ; FILE REFERENCE: JHU1120-17  
 ; CURRENT APPLICATION NUMBER: US/09/872,856  
 ; CURRENT FILING DATE: 2001-06-01  
 ; PRIOR APPLICATION NUMBER: US 09/124,180  
 ; PRIOR FILING DATE: 1998-07-28  
 ; PRIOR APPLICATION NUMBER: US 09/019,070  
 ; PRIOR FILING DATE: 1998-02-05  
 ; PRIOR APPLICATION NUMBER: US 08/862,445  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: US 08/847,910  
 ; PRIOR FILING DATE: 1997-04-28  
 ; PRIOR APPLICATION NUMBER: US 08/795,071  
 ; PRIOR FILING DATE: 1997-02-05  
 ; PRIOR APPLICATION NUMBER: US 08/525,596  
 ; PRIOR FILING DATE: 1995-10-25  
 ; PRIOR APPLICATION NUMBER: PCT/US 94/03019  
 ; PRIOR FILING DATE: 1994-03-18  
 ; PRIOR APPLICATION NUMBER: US 08/033,923  
 ; PRIOR FILING DATE: 1993-03-19  
 ; NUMBER OF SEQ ID NOS: 53  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 47  
 ; LENGTH: 115  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-872-856-47

Query Match 93.4%; Score 640; DB 11; Length 115;  
 Best Local Similarity 99.1%; Pred. No. 8e-60;  
 Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : |||||  
 Db 3 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 62

Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 |||||  
 Db 63 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 115

RESULT 7

US-10-335-483-29

; Sequence 29, Application US/10335483

; Publication No. US20030120058A1

; GENERAL INFORMATION:

; APPLICANT: Huynh, Thanh

; Lee, Se-Jin

; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-8

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CA

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/335,483

; FILING DATE: 31-Dec-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/177,860

; FILING DATE: <Unknown>

; APPLICATION NUMBER: 08/525,596

; FILING DATE: 19-SEP-1995

; APPLICATION NUMBER: PCT/US94/07762

; FILING DATE: 08-JUL-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Wetherell, Jr., Ph.D, John R.

; REGISTRATION NUMBER: 31,678

; REFERENCE/DOCKET NUMBER: 07265/075001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-678-5070

; TELEFAX: 619-678-5099

; INFORMATION FOR SEQ ID NO: 29:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 115 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; IMMEDIATE SOURCE:

; CLONE: TGF-beta-1

; FEATURE:

; NAME/KEY: Protein

; LOCATION: 1..115

; SEQUENCE DESCRIPTION: SEQ ID NO: 29:

US-10-335-483-29

Query Match 93.4%; Score 640; DB 15; Length 115;

Best Local Similarity 99.1%; Pred. No. 8e-60;

Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      :|
Db      3 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 62

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      ||
Db      63 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 115

```

RESULT 8

US-10-002-278-8

```

; Sequence 8, Application US/10002278
; Publication No. US20020132334A1
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M.
; APPLICANT: Basler, Konard
; APPLICANT: Yamada, Toshiya
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF DORSALIN-1
; FILE REFERENCE: 0575/40314-A
; CURRENT APPLICATION NUMBER: US/10/002,278
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: COOH-terminus of TGF-Beta 1
; NAME/KEY: DOMAIN
; LOCATION: (1)..(139)
; OTHER INFORMATION:

```

US-10-002-278-8

```

Query Match          93.4%; Score 640; DB 14; Length 139;
Best Local Similarity 99.1%; Pred. No. 9.9e-60;
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      :|
Db      27 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 86

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      ||
Db      87 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 139

```

RESULT 9

US-09-214-592-25

```

; Sequence 25, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
; APPLICANT: Yamasaki, CMotoo
; APPLICANT: Shibata, CKenji
; APPLICANT: Sato, CYasufumi

```

```

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 25
; LENGTH: 315
; TYPE: PRT
; ORGANISM: bovine
US-09-214-592-25

```

: TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

FILE REFERENCE: 11060

; CURRENT APPLICATION NUMBER: US/09/214,592A

; CURRENT FILING DATE: 1999-01-18

; NUMBER OF SEO ID NOS: 34

; SOFTWARE:

; SEQ ID NO 25

LENGTH: 315

; TYPE: PRT

ORGANISM: bovine

US-09-214-592-25

Query Match 93.4%; Score 640; DB 11; Length 315;  
Best Local Similarity 99.1%; Pred. No. 2.4e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
:|||||  
Db 203 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 262

QY           68 KVLALYNQHNP GASAAPCCVPQA LEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db         263 KVLALYNQHNP GASAAPCCVPQA LEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS 315

RESULT 10

US-09-756-283A-23

; Sequence 23, Application US/09756283A

; Patent No. US20020151478A1

; GENERAL INFORMATION:

APPLICANT: Chernajovsky, Yuti

; APPLICANT: Dreja, Hanna Stina

; APPLICANT: Adams, Gillian

TITLE OF INVENTION: Latent Fusion Protein

; FILE REFERENCE: 0623.1000000

CURRENT APPLICATION NUMBER: US/09/756,283A

; CURRENT FILING DATE: 2001-01-09

; NUMBER OF SEQ ID NOS: 100

```

; SOFTWARE: PatentIn version 3.0

```

; SEO ID NO 23

; LENGTH: 390

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-756-283A-23

Query Match 93.4%; Score 640; DB 10; Length 390;  
Best Local Similarity 99.1%; Pred. No. 3e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
: |||||  
Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

QY           68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 338 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 11

US-09-214-592-26

; Sequence 26, Application US/09214592A  
; Publication No. US20030027218A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamasaki,CMotoo  
; APPLICANT: Shibata,CKenji  
; APPLICANT: Sato,CYasufumi  
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND  
METHOD  
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS  
; FILE REFERENCE: 11060  
; CURRENT APPLICATION NUMBER: US/09/214,592A  
; CURRENT FILING DATE: 1999-01-18  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE:  
; SEQ ID NO 26  
; LENGTH: 390  
; TYPE: PRT  
; ORGANISM: porcine  
US-09-214-592-26

Query Match 93.4%; Score 640; DB 11; Length 390;  
Best Local Similarity 99.1%; Pred. No. 3e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
  
Qy 68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 338 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 12

US-09-214-592-28

; Sequence 28, Application US/09214592A  
; Publication No. US20030027218A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamasaki,CMotoo  
; APPLICANT: Shibata,CKenji  
; APPLICANT: Sato,CYasufumi  
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND  
METHOD  
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS  
; FILE REFERENCE: 11060  
; CURRENT APPLICATION NUMBER: US/09/214,592A  
; CURRENT FILING DATE: 1999-01-18  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE:  
; SEQ ID NO 28  
; LENGTH: 390  
; TYPE: PRT

; ORGANISM: canine  
US-09-214-592-28

Query Match 93.4%; Score 640; DB 11; Length 390;  
Best Local Similarity 99.1%; Pred. No. 3e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 13

US-09-214-592-29

; Sequence 29, Application US/09214592A  
; Publication No. US20030027218A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamasaki,CMotoo  
; APPLICANT: Shibata,CKenji  
; APPLICANT: Sato,CYasufumi  
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND  
METHOD  
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS  
; FILE REFERENCE: 11060  
; CURRENT APPLICATION NUMBER: US/09/214,592A  
; CURRENT FILING DATE: 1999-01-18  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE:  
; SEQ ID NO 29  
; LENGTH: 390  
; TYPE: PRT  
; ORGANISM: ovine  
US-09-214-592-29

Query Match 93.4%; Score 640; DB 11; Length 390;  
Best Local Similarity 99.1%; Pred. No. 3e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 14

US-09-214-592-33

; Sequence 33, Application US/09214592A  
; Publication No. US20030027218A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamasaki,CMotoo



Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 |||

Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

Search completed: October 28, 2003, 09:27:59  
 Job time : 7.28061 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 00:37:16 ; Search time 3.96533 Seconds  
 (without alignments)  
 2910.285 Million cell updates/sec

Title: US-10-017-372E-11  
 Perfect score: 685  
 Sequence: 1 DYKDDDDKALDTNYCFSSTE.....GRKPKVEQLSNMIVRSCKCS 120

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : PIR\_76:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	Score	% Query Match	Length	DB	ID	Description
1	640	93.4	315	2	A40057	transforming growth
2	640	93.4	390	1	WFHU2	transforming growth
3	640	93.4	390	2	JC4023	transforming growth
4	640	93.4	390	2	A26960	transforming growth



5	640	93.4	390	2	A27512	transforming growth
6	640	93.4	390	2	I46463	transforming growth
7	640	93.4	391	2	S01413	transforming growth
8	637	93.0	390	1	WFMS2	transforming growth
9	637	93.0	390	2	S10219	transforming growth
10	625	91.2	130	2	I48196	transforming growth
11	542	79.1	373	2	A41918	transforming growth
12	505	73.7	382	2	B61036	transforming growth
13	502	73.3	410	2	A41397	transforming growth
14	502	73.3	412	2	A36169	transforming growth
15	501	73.1	410	2	A55706	transforming growth
16	499	72.8	412	2	A34939	transforming growth
17	495	72.3	409	2	S01825	transforming growth
18	491	71.7	412	2	A39489	transforming growth
19	484	70.7	414	1	WFMKB2	transforming growth
20	484	70.7	414	2	A31249	transforming growth
21	484	70.7	442	2	B31249	transforming growth
22	482	70.4	112	2	A61439	transforming growth
23	482	70.4	414	1	WFMSB2	transforming growth
24	480	70.1	413	1	WFXLB2	transforming growth
25	204	29.8	425	2	I47072	inhibin beta-A cha
26	200	29.2	424	1	WFPGBA	inhibin beta-A cha
27	200	29.2	424	1	S31440	inhibin beta-A cha
28	200	29.2	425	1	S50898	inhibin beta-A cha
29	200	29.2	426	1	B24248	inhibin beta-A cha
30	196	28.6	424	1	B40905	inhibin beta-A cha
31	185	27.0	350	2	JC5241	activin beta E cha
32	181	26.4	413	2	JC4862	activin beta-A cha
33	179	26.1	115	2	PN0504	activin beta A cha
34	177	25.8	455	2	A43918	TGF-beta-related p
35	174	25.4	207	2	S37618	vgr protein - rat
36	174	25.4	367	2	JC4151	activin beta D cha
37	174	25.4	513	1	BMHU6	bone morphogenetic
38	173	25.3	510	2	A54798	Vg-1-related prote
39	171	25.0	43	2	B26356	transforming growth
40	169	24.7	352	2	JC2466	inhibin beta-C cha
41	169	24.7	372	2	C39364	GDF-1 embryonic gr
42	168.5	24.6	360	2	A29619	Vgl embryonic grow
43	167.5	24.5	370	2	I51199	activin beta B sub
44	167.5	24.5	430	2	JQ1184	osteogenic protein
45	166.5	24.3	313	2	I51284	bone morphogenetic

#### ALIGNMENTS

##### RESULT 1

A40057

transforming growth factor beta-1 precursor - bovine (fragment)

N;Alternate names: beta-TGF; cartilage-inducing factor-A; EGF-dependent TGF or dEGF; MGF-b; milk growth factor b; TGF-type II

C;Species: Bos primigenius taurus (cattle)

C;Date: 28-Feb-1992 #sequence\_revision 28-Feb-1992 #text\_change 16-Jul-1999

C;Accession: A40057; A42320; A05284; A24322; B61439

R;Van Obberghen-Schilling, E.; Kondaiah, P.; Ludwig, R.L.; Sporn, M.B.; Baker, C.C.

Mol. Endocrinol. 1, 693-698, 1987

A;Title: Complementary deoxyribonucleic acid cloning of bovine transforming growth factor-beta1.  
A;Reference number: A40057; MUID:91042552; PMID:3153459  
A;Accession: A40057  
A;Molecule type: mRNA  
A;Residues: 1-315 <VAN>  
A;Cross-references: GB:M36271; NID:g163747; PIDN:AAA30778.1; PID:g163748  
R;Ogawa, Y.; Schmidt, D.K.; Dasch, J.R.; Chang, R.J.; Glaser, C.B.  
J. Biol. Chem. 267, 2325-2328, 1992  
A;Title: Purification and characterization of transforming growth factor-beta2.3 and -beta1.2 heterodimers from bovine bone.  
A;Reference number: A42320; MUID:92129307; PMID:1733936  
A;Accession: A42320  
A;Molecule type: protein  
A;Residues: 204-209,'X',211-217 <OGA>  
R;Roberts, A.B.; Anzano, M.A.; Meyers, C.A.; Wideman, J.; Blacher, R.; Pan, Y.C.E.; Stein, S.; Lehrman, S.R.; Smith, J.M.; Lamb, L.C.; Sporn, M.B.  
Biochemistry 22, 5692-5698, 1983  
A;Title: Purification and properties of a type beta transforming growth factor from bovine kidney.  
A;Reference number: A05284; MUID:84104793; PMID:6607069  
A;Accession: A05284  
A;Molecule type: protein  
A;Residues: 204-218 <ROB>  
R;Seyedin, S.M.; Thompson, A.Y.; Bentz, H.; Rosen, D.M.; McPherson, J.M.; Conti, A.; Siegel, N.R.; Galluppi, G.R.; Piez, K.A.  
J. Biol. Chem. 261, 5693-5695, 1986  
A;Title: Cartilage-inducing factor-A. Apparent identity to transforming growth factor-beta.  
A;Reference number: A24322; MUID:86195954; PMID:3754555  
A;Accession: A24322  
A;Molecule type: protein  
A;Residues: 204-233 <SEY>  
R;Jin, Y.; Cox, D.A.; Knecht, R.; Raschdorf, F.; Cerletti, N.  
J. Protein Chem. 10, 565-575, 1991  
A;Title: Separation, purification, and sequence identification of TGF-beta1 and TGF-beta2 from bovine milk.  
A;Reference number: A61439; MUID:92189724; PMID:1799413  
A;Accession: B61439  
A;Molecule type: protein  
A;Residues: 204-209,'X',211-217,'XX',220-232 <JIN>  
C;Comment: This polypeptide is composed of two polypeptide chains cross-linked by disulfide bonds. It has been found in neoplastic and non-neoplastic tissues.  
C;Comment: Type II TGF does not bind to the EGF receptor and lacks intrinsic mitogenic activity, but in soft agar, it reacts synergistically with either type I TFG or EGF, and induces cell proliferation. Cells grown in monolayer do not respond in a similar manner to these growth factors, but morphologically do acquire a transformed phenotype.  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; heterodimer  
F;204-315/Product: transforming growth factor beta-1 #status experimental <MAT>  
F;7,61,101/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.4%; Score 640; DB 2; Length 315;  
Best Local Similarity 99.1%; Pred. No. 9.4e-60;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      203 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 262

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      263 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 315

```

## RESULT 2

WFHU2

transforming growth factor beta-1 precursor [validated] - human

N;Alternate names: growth-inhibitory factor; TGF type 2; TGF-beta

C;Species: Homo sapiens (man)

C;Date: 28-Feb-1986 #sequence\_revision 19-Oct-1995 #text\_change 08-Dec-2000

C;Accession: A27513; A01395; A22290; I59664; S53444

R;Derynck, R.; Rhee, L.; Chen, E.Y.; Van Tilburg, A.

Nucleic Acids Res. 15, 3188-3189, 1987

A;Title: Intron-exon structure of the human transforming growth factor-beta precursor gene.

A;Reference number: A27513; MUID:87174845; PMID:3470709

A;Accession: A27513

A;Molecule type: DNA

A;Residues: 1-390 <DER>

A;Cross-references: GB:X05839; GB:Y00112; NID:g37097; PIDN:CAA29283.1;

PID:g1212989

R;Derynck, R.; Jarrett, J.A.; Chen, E.Y.; Eaton, D.H.; Bell, J.R.; Assoian, R.K.; Roberts, A.B.; Sporn, M.B.; Goeddel, D.V.

Nature 316, 701-705, 1985

A;Title: Human transforming growth factor-beta complementary DNA sequence and expression in normal and transformed cells.

A;Reference number: A01395; MUID:85296301; PMID:3861940

A;Accession: A01395

A;Molecule type: mRNA

A;Residues: 1-9, 'P', 11-24, 'P', 26-159, 'R', 160-390 <DE2>

A;Cross-references: GB:X02812; GB:J05114; NID:g37092; PIDN:CAA26580.1;

PID:g37093

A;Note: the authors suggest that residues 8-23 could represent the hydrophobic core of an amino-terminal signal peptide

R;Massague, J.; Like, B.

J. Biol. Chem. 260, 2636-2645, 1985

A;Title: Cellular receptors for type beta transforming growth factor. Ligand binding and affinity labeling in human and rodent cell lines.

A;Reference number: A22290; MUID:85131019; PMID:2982829

A;Accession: A22290

A;Molecule type: protein

A;Residues: 279-295, 'XX', 298-301 <MAS>

R;Urushizaki, Y.; Niitsu, Y.; Terui, T.; Koshida, Y.; Mahara, K.; Kohgo, Y.; Urushizaki, I.; Takahashi, Y.; Ito, H.

Tumor Res. 22, 41-55, 1987

A;Title: Cloning and expression of the gene for human transforming growth factor-beta in Escherichia coli.

A;Reference number: I59664

A;Accession: I59664

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 279-390 <RES>

A;Cross-references: GB:M38449; NID:g339557; PIDN:AAA36735.1; PID:g339558  
R;Stam, K.; Stewart, A.A.; Qu, G.Y.; Iwata, K.K.; Fenyoe, D.; Chait, B.T.;  
Marshak, D.R.; Haley, J.D.  
Biochem. J. 305, 87-92, 1995  
A;Title: Physical and biological characterization of a growth-inhibitory  
activity purified from the neuroepithelioma cell line A673.  
A;Reference number: S53444; MUID:95126934; PMID:7826358  
A;Accession: S53444  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 279-297 <STA>  
C;Comment: The mature protein is the carboxyl-terminal segment of a precursor  
polypeptide; the active molecule is a dimer of identical polypeptide chains  
linked by an interchain disulfide bond.  
C;Genetics:  
A;Gene: GDB:TGFB1; TGFB  
A;Cross-references: GDB:120729; OMIM:190180  
A;Map position: 19q13.2-19q13.2  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; homodimer; mitogen; transformation  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-278/Domain: propeptide #status predicted <PRO>  
F;244-246/Region: cell attachment (R-G-D) motif  
F;279-390/Product: transforming growth factor beta-1 #status experimental <MAT>  
F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.4%; Score 640; DB 1; Length 390;  
Best Local Similarity 99.1%; Pred. No. 1.2e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      :|
Db      278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||
Db      338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

# RESULT 3

JC4023

transforming growth factor beta-1 - dog

C;Species: Canis lupus familiaris (dog)

C;Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 24-Nov-1999

C;Accession: JC4023

R;Manning, A.M.; Auchampach, J.A.; Drong, R.F.; Slightom, J.L.

Gene 155, 307-308, 1995

A;Title: Cloning of a canine cDNA homologous to the human transforming growth  
factor-beta 1-encoding gene.

A;Reference number: JC4023; MUID:95237630; PMID:7721110

A;Accession: JC4023

A;Molecule type: mRNA

A;Residues: 1-390 <MAN>

A;Cross-references: GB:L34956; NID:g516071; PIDN:AAA51458.1; PID:g516072

C;Comment: This factor plays a multifunctional role as a regulator of mammalian  
cell growth and as a modulator of immune responses.

C;Genetics:

A;Gene: tgf-betal  
C;Superfamily: inhibin  
C;Keywords: growth factor; transforming protein  
F;288-390/Product: transforming growth factor beta 1 #status predicted <MAT>

Query Match 93.4%; Score 640; DB 2; Length 390;  
Best Local Similarity 99.1%; Pred. No. 1.2e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      :|
Db      278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||
Db      338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

#### RESULT 4

A26960

transforming growth factor beta-1 precursor - green monkey

C;Species: Cercopithecus aethiops (green monkey, grivet)

C;Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 24-Nov-1999

C;Accession: A26960

R;Sharples, K.; Plowman, G.D.; Rose, T.M.; Twardzik, D.R.; Purchio, A.F.

DNA 6, 239-244, 1987

A;Title: Cloning and sequence analysis of simian transforming growth factor-beta cDNA.

A;Reference number: A26960; MUID:87246074; PMID:3474130

A;Accession: A26960

A;Molecule type: mRNA

A;Residues: 1-390 <SHA>

A;Cross-references: GB:M16658; NID:g176552; PIDN:AAA35369.1; PID:g176553

C;Superfamily: inhibin

C;Keywords: growth factor

F;1-16/Domain: signal sequence #status predicted <SIG>

F;17-390/Product: transforming growth factor beta #status predicted <MAT>

Query Match 93.4%; Score 640; DB 2; Length 390;  
Best Local Similarity 99.1%; Pred. No. 1.2e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      :|
Db      278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||
Db      338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

#### RESULT 5

A27512

transforming growth factor beta-1 precursor - pig

N;Alternate names: TGF-beta

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 05-Jun-1988 #sequence\_revision 05-Jun-1988 #text\_change 16-Jul-1999

C;Accession: A27512; A26356; I46657  
R;Derynck, R.; Rhee, L.  
Nucleic Acids Res. 15, 3187, 1987  
A;Title: Sequence of the porcine transforming growth factor-beta precursor.  
A;Reference number: A27512; MUID:87174844; PMID:3470708  
A;Accession: A27512  
A;Molecule type: mRNA  
A;Residues: 1-390 <DER>  
R;Cheifetz, S.; Weatherbee, J.A.; Tsang, M.L.S.; Anderson, J.K.; Mole, J.E.;  
Lucas, R.; Massague, J.  
Cell 48, 409-415, 1987  
A;Title: The transforming growth factor-beta system, a complex pattern of cross-  
reactive ligands and receptors.  
A;Reference number: A90890; MUID:87102890; PMID:2879635  
A;Accession: A26356  
A;Molecule type: protein  
A;Residues: 279-322 <CHE>  
R;Kondaiah, P.; Van Obberghen-Schilling, E.; Ludwig, R.L.; Dhar, R.; Sporn,  
M.B.; Roberts, A.B.  
J. Biol. Chem. 263, 18313-18317, 1988  
A;Title: cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.  
Evidence for alternate splicing and polyadenylation.  
A;Reference number: I46657; MUID:89054010; PMID:2461367  
A;Accession: I46657  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-390 <KON>  
A;Cross-references: GB:M23703; NID:g755044; PIDN:AAA64616.1; PID:g755045  
C;Genetics:  
A;Gene: TGFB; TGF-beta-1  
C;Superfamily: inhibin  
C;Keywords: growth factor

Query Match 93.4%; Score 640; DB 2; Length 390;  
Best Local Similarity 99.1%; Pred. No. 1.2e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      : ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      : ||||||||||||||||||||||||||||||||||||||||||||
Db      338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 6  
I46463  
transforming growth factor beta-1 - sheep  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 19-Dec-1997 #sequence\_revision 19-Dec-1997 #text\_change 24-Nov-1999  
C;Accession: I46463; S45115  
R;Woodall, C.J.; McLaren, L.J.; Watt, N.J.  
Gene 150, 371-373, 1994  
A;Title: Sequence and chromosomal localisation of the gene encoding ovine latent  
transforming growth factor-beta 1.  
A;Reference number: I46463; MUID:95121932; PMID:7821809

A;Accession: I46463  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-390 <WOO>  
A;Cross-references: EMBL:X76916; NID:g496648; PIDN:CAA54242.1; PID:g496649  
A;Note: submitted to the EMBL Data Library, December 1993  
C;Superfamily: inhibin

Query Match 93.4%; Score 640; DB 2; Length 390;  
Best Local Similarity 99.1%; Pred. No. 1.2e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy     68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390
```

#### RESULT 7

S01413  
transforming growth factor beta-1 precursor - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 24-Nov-1999  
C;Accession: S01413  
R;Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B.  
Nucleic Acids Res. 16, 8730, 1988  
A;Title: Nucleotide sequence of chicken transforming growth factor-beta 1 (TGF-beta 1).  
A;Reference number: S01413; MUID:88335639; PMID:3166520  
A;Accession: S01413  
A;Molecule type: DNA  
A;Residues: 1-391 <JAK>  
A;Cross-references: EMBL:X12373; NID:g63808; PIDN:CAA30933.1; PID:g63809  
C;Superfamily: inhibin  
C;Keywords: growth factor

Query Match 93.4%; Score 640; DB 2; Length 391;  
Best Local Similarity 99.1%; Pred. No. 1.2e-59;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     279 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 338

Qy     68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    339 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 391
```

#### RESULT 8

WFMS2  
transforming growth factor beta-1 precursor - mouse  
N;Alternate names: TGF type 2; TGF-beta  
C;Species: Mus musculus (house mouse)

C;Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 24-Nov-1999  
 C;Accession: A01396  
 R;Derynck, R.; Jarrett, J.A.; Chen, E.Y.; Goeddel, D.V.  
 J. Biol. Chem. 261, 4377-4379, 1986  
 A;Title: The murine transforming growth factor-beta precursor.  
 A;Reference number: A01396; MUID:86168129; PMID:3007454  
 A;Accession: A01396  
 A;Molecule type: mRNA  
 A;Residues: 1-390 <DER>  
 A;Cross-references: GB:M13177; NID:g201952; PIDN:AAA40423.1; PID:g201953  
 A;Note: the authors suggest that residues 8-23 could represent the hydrophobic core of an amino-terminal signal peptide  
 C;Comment: The mature protein is the carboxyl-terminal segment of a precursor polypeptide; the active molecule is a dimer of identical polypeptide chains linked by an interchain disulfide bond.  
 C;Superfamily: inhibin  
 C;Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen; transformation  
 F;1-23/Domain: signal sequence #status predicted <SIG>  
 F;24-278/Domain: propeptide #status predicted <PRO>  
 F;244-246/Region: cell attachment (R-G-D) motif  
 F;279-390/Product: transforming growth factor beta-1 #status predicted <MAT>  
 F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.0%; Score 637; DB 1; Length 390;  
 Best Local Similarity 98.2%; Pred. No. 2.4e-59;  
 Matches 111; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      :|
Db      278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy      68 KVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      |||
Db      338 KVLALYNQHNPASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

#### RESULT 9

S10219  
 transforming growth factor beta-1 precursor - rat  
 N;Alternate names: TGF type 2; TGF-beta  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 24-Nov-1999  
 C;Accession: S10219; PT0023; S02267  
 R;Qian, S.W.; Kondaiah, P.; Roberts, A.B.; Sporn, M.B.  
 Nucleic Acids Res. 18, 3059, 1990  
 A;Title: cDNA cloning by PCR of rat transforming growth factor beta-1.  
 A;Reference number: S10219; MUID:90272425; PMID:2349108  
 A;Accession: S10219  
 A;Molecule type: mRNA  
 A;Residues: 1-390 <QIA>  
 A;Cross-references: EMBL:X52498; NID:g57341; PIDN:CAA36741.1; PID:g57342  
 R;Okada, F.; Yamaguchi, K.; Ichihara, A.; Nakamura, T.  
 J. Biochem. 106, 304-310, 1989  
 A;Title: Purification and structural analysis of a latent form of transforming growth factor-beta from rat platelets.  
 A;Reference number: PT0023; MUID:90036779; PMID:2478527



A;Accession: PT0023  
A;Molecule type: protein  
A;Residues: 30-32,'X',34-38,'Q',40-42,'X',44 <OKA>  
R;Okada, F.; Yamaguchi, K.; Ichihara, A.; Nakamura, T.  
FEBS Lett. 242, 240-244, 1989  
A;Title: One of two subunits of masking protein in latent TGF-beta is a part of pro-TGF-beta.  
A;Reference number: S02267; MUID:89121078; PMID:2914605  
A;Accession: S02267  
A;Molecule type: protein  
A;Residues: 30-32,'X',34-38,'Q',40-42,'X',44 <OK2>  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; integrin binding  
F;1-29/Domain: signal sequence #status predicted <SIG>  
F;30-278/Domain: propeptide #status experimental <PRO>  
F;244-246/Region: cell attachment (R-G-D) motif  
F;279-390/Product: transforming growth factor beta-1 #status predicted <MAT>  
F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.0%; Score 637; DB 2; Length 390;  
Best Local Similarity 98.2%; Pred. No. 2.4e-59;  
Matches 111; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      :|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy      68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      |||||:|||||||||||||||||||||||||||||||||
Db      338 KVLALYNQHNP GASASPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390
```

# RESULT 10

I48196  
transforming growth factor beta-1 precursor - golden hamster (fragment)  
C;Species: Mesocricetus auratus (golden hamster)  
C;Date: 02-Jul-1996 #sequence\_revision 04-Oct-1996 #text\_change 24-Nov-1999  
C;Accession: I48196  
R;Wong, D.T.; Donoff, R.B.; Yang, J.; Song, B.Z.; Matossian, K.; Nagura, N.;  
Elovic, A.; McBride, J.; Gallagher, G.; Todd, R.; Chiang, T.; Chou, L.S.S.;  
Yung, C.M.; Galli, S.J.; Weller, P.F.  
Am. J. Pathol. 143, 130-142, 1993  
A;Title: Sequential expression of transforming growth factors alpha and beta 1  
by eosinophils during cutaneous wound healing in the hamster.  
A;Reference number: I48196; MUID:93304479; PMID:8317544  
A;Accession: I48196  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-130 <RES>  
A;Cross-references: EMBL:X60296; NID:g396177; PIDN:CAA42838.1; PID:g396178  
C;Superfamily: inhibin

Query Match 91.2%; Score 625; DB 2; Length 130;  
Best Local Similarity 97.3%; Pred. No. 1.4e-58;  
Matches 110; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
```

```

      :|||||||||||||||||||||||||||||||||||||||||||||
Db      18 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 77

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||||||||||||||||||||||||||||||||||||||||
Db      78 KVLALYNQHNPASAGPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSYKCS 130

```

RESULT 11

A41918  
transforming growth factor beta-4 precursor - chicken (fragment)  
N;Alternate names: TGF-beta 4  
C;Species: Gallus gallus (chicken)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
C;Accession: A41918; A34941; S03110  
R;Burt, D.W.; Jakowlew, S.B.  
Mol. Endocrinol. 6, 989-992, 1992  
A;Title: Correction: a new interpretation of a chicken transforming growth factor-beta 4 complementary DNA.  
A;Reference number: A41918; MUID:92357039; PMID:1353860  
A;Accession: A41918  
A;Molecule type: mRNA  
A;Residues: 1-373 <BUR>  
A;Cross-references: GB:M31160; GB:X08012; GB:S41706; NID:g1262437; PIDN:AAB05637.1; PID:g1262438  
A;Note: sequence extracted from NCBI backbone (NCBIN:110186, NCBIIP:110187)  
A;Note: this report corrects and reinterprets the sequence from reference A34941  
R;Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B.  
Mol. Endocrinol. 2, 1186-1195, 1988  
A;Title: Complementary deoxyribonucleic acid cloning of a messenger ribonucleic acid encoding transforming growth factor beta-4 from chicken embryo chondrocytes.  
A;Reference number: A34941; MUID:89112198; PMID:2464131  
A;Accession: A34941  
A;Molecule type: mRNA  
A;Residues: 'MDPMSIGPKSCGGSPWRPPGTAPWSIGSRRATASSSCSTSSRVRAEVGGRAL',122-209,'D',211-373 <JAK>  
A;Cross-references: EMBL:X08012  
A;Note: this sequence has been corrected in A41918  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor  
F;1/Domain: signal sequence (fragment) #status predicted <SIG>  
F;223-225/Region: cell attachment (R-G-D) motif  
F;260-373/Product: transforming growth factor beta-4 #status predicted <MAT>  
F;54,109,153/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Query Match          79.1%;  Score 542;  DB 2;  Length 373;
Best Local Similarity 82.3%;  Pred. No. 2.3e-49;
Matches 93;  Conservative 9;  Mismatches 9;  Indels 2;  Gaps 1;

```

```

Qy      10 LDTNYCF--SSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      |||:||||  : ||||||| ||||||||| ||||||||| |||:||||||| ||||:
Db      261 LDTDYCFGPGTDEKNCCVRPLYIDFRKDLQWKWIHEPKGYMANFCMGPCPYIWSADTQYT 320

Qy      68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||||||||||||||||||||||| :|||:||||| :||||||:|:||||
Db      321 KVLALYNQHNPASAAAPCCVPQTLDPLPIIYYVGRNVRVEQLSNMIVRVACKCS 373

```

B61036

Query Match 73.7%; Score 505; DB 2; Length 382;  
Best Local Similarity 74.3%; Pred. No. 1.9e-45;  
Matches 84; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
: : ||| : ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||  
Db 270 RGVGQEYCFGNGPNCCVKPLYINFRKDLGWKWIHEPKGYEANYCLGNCPYIWSMDTQYS 329

Qy 68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
|||: |||: |||| : |||| ||||: |||| ||||: ||||: ||||: ||||: ||||: ||||  
Db 330 KVLSLYNNNPGASISPCCVDPVLEPLPIIYYVGRTAKVEOLSNMIVRSCNCS 382

## A41397

transforming growth factor beta-3 precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 03-Apr-1992 #sequence\_revision 03-Apr-1992 #text\_change 16-Jul-1999  
C;Accession: A41397; A61039; A61225  
R;Miller, D.A.; Lee, A.; Matsui, Y.; Chen, E.Y.; Moses, H.L.; Derynck, R.  
Mol. Endocrinol. 3, 1926-1934, 1989  
A;Title: Complementary DNA cloning of the murine transforming growth factor-  
beta3 (TGFbeta3) precursor and the comparative expression of TGFbeta3 and  
TGFbetal messenger RNA in murine embryos and adult tissues.  
A;Reference number: A41397; MUID:90190650; PMID:2628730

A;Accession: A41397  
A;Molecule type: mRNA  
A;Residues: 1-410 <MIL>  
A;Cross-references: GB:M32745; NID:g201949; PIDN:AAA40422.1; PID:g201950  
R;Denhez, F.; Lafyatis, R.; Kondaiah, P.; Roberts, A.B.; Sporn, M.B.  
Growth Factors 3, 139-146, 1990  
A;Title: Cloning by polymerase chain reaction of a new mouse TGF-beta, mTGF-beta3.  
A;Reference number: A61039; MUID:91000714; PMID:2206556  
A;Accession: A61039  
A;Molecule type: mRNA  
A;Residues: 1-410 <DEN>  
R;Watrin, F.; Scotto, L.; Assoian, R.K.; Wolgemuth, D.J.  
Cell Growth Differ. 2, 77-83, 1991  
A;Title: Cell lineage specificity of expression of the murine transforming growth factor beta-3 and transforming growth factor beta-1 genes.  
A;Reference number: A61225; MUID:91299576; PMID:2069871  
A;Accession: A61225  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 285-410 <WAT>  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; growth regulation  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-298/Domain: propeptide #status predicted <PRO>  
F;259-261/Region: cell attachment (R-G-D) motif  
F;299-410/Product: transforming growth factor beta-3 #status predicted <MAT>  
F;72,133,140/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 73.3%; Score 502; DB 2; Length 410;  
Best Local Similarity 76.1%; Pred. No. 4.2e-45;  
Matches 86; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

```

Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      :| | | | | | | | : | : | | | | | | | | | | | | | | | | | | | | | | | |
Db      298 RALDTNYCFRNLEENCCVRPLYIDFRQDLGWKQVHEPKGYANFCSGPCPYLRSADTTHS 357

Qy      68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      358 TVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCKCS 410

```

RESULT 14  
A36169  
transforming growth factor beta-3 precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 20-Jun-2000  
C;Accession: A36169; A41262; S01824  
R;ten Dijke, P.; Hansen, P.; Iwata, K.K.; Pieler, C.; Foulkes, J.G.  
Proc. Natl. Acad. Sci. U.S.A. 85, 4715-4719, 1988  
A;Title: Identification of another member of the transforming growth factor type beta gene family.  
A;Reference number: A36169; MUID:88263019; PMID:3164476  
A;Accession: A36169  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-412 <TEN>





Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	640	93.4	315	1	TGF1_BOVIN	P18341 bos taurus
2	640	93.4	390	1	TGF1_CANFA	P54831 canis famil
3	640	93.4	390	1	TGF1_CERAE	P09533 cercopithec
4	640	93.4	390	1	TGF1_HUMAN	P01137 homo sapien
5	640	93.4	390	1	TGF1_PIG	P07200 sus scrofa
6	640	93.4	390	1	TGF1_SHEEP	P50414 ovis aries
7	637	93.0	390	1	TGF1_MOUSE	P04202 mus musculu
8	637	93.0	390	1	TGF1_RAT	P17246 rattus norv
9	636	92.8	390	1	TGF1_HORSE	O19011 equus cabal
10	628	91.7	390	1	TGF1_CAVPO	Q9z1y6 cavia porce
11	542	79.1	373	1	TGF1_CHICK	P09531 gallus gall
12	505	73.7	382	1	TGF1_XENLA	P16176 xenopus lae
13	502	73.3	410	1	TGF3_MOUSE	P17125 mus musculu
14	502	73.3	412	1	TGF3_HUMAN	P10600 homo sapien
15	501	73.1	412	1	TGF3_RAT	Q07258 rattus norv
16	495	72.3	409	1	TGF3_PIG	P15203 sus scrofa
17	491	71.7	412	1	TGF2_CHICK	P30371 gallus gall
18	484	70.7	414	1	TGF2_HUMAN	P08112 homo sapien
19	484	70.7	435	1	TGF2_PIG	P09858 sus scrofa
20	482	70.4	112	1	TGF2_BOVIN	P21214 bos taurus
21	482	70.4	414	1	TGF2_MOUSE	P27090 mus musculu
22	482	70.4	442	1	TGF2_RAT	Q07257 rattus norv
23	480	70.1	412	1	TGF3_CHICK	P16047 gallus gall
24	480	70.1	413	1	TGF2_XENLA	P17247 xenopus lae
25	204	29.8	425	1	IHBA_SHEEP	P43032 ovis aries
26	204	29.8	426	1	IHBA_HORSE	P55102 equus cabal
27	202.5	29.6	374	1	GDF8_BRARE	O42222 brachydanio
28	200	29.2	424	1	IHBA_MOUSE	Q04998 mus musculu
29	200	29.2	424	1	IHBA_PIG	P03970 sus scrofa
30	200	29.2	424	1	IHBA_RAT	P18331 rattus norv
31	200	29.2	425	1	IHBA_BOVIN	P07995 bos taurus
32	200	29.2	426	1	IHBA_HUMAN	P08476 homo sapien
33	199.5	29.1	375	1	GDF8_PAPHA	O18828 papio hamad
34	196.5	28.7	375	1	GDF8_SHEEP	O18830 ovis aries
35	195.5	28.5	375	1	GDF8_CHICK	O42220 gallus gall
36	195.5	28.5	375	1	GDF8_HUMAN	O14793 homo sapien
37	195.5	28.5	375	1	GDF8_MELGA	O42221 meleagris g
38	195.5	28.5	375	1	GDF8_PIG	O18831 sus scrofa
39	195.5	28.5	376	1	GDF8_MOUSE	O08689 mus musculu
40	195.5	28.5	376	1	GDF8_RAT	O35312 rattus norv
41	191	27.9	424	1	IHBA_CHICK	P27092 gallus gall
42	190.5	27.8	355	1	DVR1_BRARE	P35621 brachydanio
43	190.5	27.8	375	1	GDF8_BOVIN	O18836 bos taurus
44	188.5	27.5	405	1	GDFB_MOUSE	Q9z1w4 mus musculu
45	188.5	27.5	407	1	GDFB_HUMAN	O95390 homo sapien

## ALIGNMENTS

## RESULT 1

TGF1\_BOVIN

ID TGF1\_BOVIN STANDARD; PRT; 315 AA.  
AC P18341;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Transforming growth factor beta 1 precursor (TGF-beta 1) (Fragment).  
GN TGFB1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91042552; PubMed=3153459;  
RA van Obberghen-Schilling E., Kondaiah P., Ludwig R.L., Sporn M.B.,  
RA Baker C.C.;  
RT "Complementary deoxyribonucleic acid cloning of bovine transforming  
RT growth factor-beta 1.";  
RL Mol. Endocrinol. 1:693-698(1987).  
RN [2]  
RP SUBUNITS.  
RC TISSUE=Bone;  
RX MEDLINE=92129307; PubMed=1733936;  
RA Ogawa Y., Schmidt D.K., Dasch J.R., Chang R.J., Glaser C.B.;  
RT "Purification and characterization of transforming growth factor-beta  
RT 2.3 and -beta 1.2 heterodimers from bovine bone.";  
RL J. Biol. Chem. 267:2325-2328(1992).  
CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
CC -!- SUBUNIT: Homodimer; disulfide-linked. Heterodimers of TGF-beta 1/2  
CC have been found in bone.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the TGF-beta family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M36271; AAA30778.1; -.  
DR PIR; A40057; A40057.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR002400; GF\_cysknot.



DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein.  
 FT NON\_TER 1 1  
 FT PROPEP <1 203  
 FT CHAIN 204 315 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 210 219 BY SIMILARITY.  
 FT DISULFID 218 281 BY SIMILARITY.  
 FT DISULFID 247 312 BY SIMILARITY.  
 FT DISULFID 251 314 BY SIMILARITY.  
 FT DISULFID 280 280 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 169 171 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 315 AA; 36269 MW; C2717A23D994E00E CRC64;

Query Match 93.4%; Score 640; DB 1; Length 315;  
 Best Local Similarity 99.1%; Pred. No. 1.2e-61;  
 Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 203 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 262  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 263 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 315

RESULT 2  
 TGF1\_CANFA  
 ID TGF1\_CANFA STANDARD; PRT; 390 AA.  
 AC P54831;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
 GN TGFB1.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Jugular vein endothelial;  
 RX MEDLINE=95237630; PubMed=7721110;  
 RA Manning A.M., Auchampach J.A., Drong R.F., Slightom J.L.;  
 RT "Cloning of a canine cDNA homologous to the human transforming growth

```

RT    factor-beta 1-encoding gene.";
RL    Gene 155:307-308(1995).
CC    -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC          PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC          TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC          HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC          ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC          A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC    -!- SUBUNIT: Homodimer; disulfide-linked.
CC    -!- SUBCELLULAR LOCATION: Secreted.
CC    -!- SIMILARITY: Belongs to the TGF-beta family.
CC    -----
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    -----
DR    EMBL; L34956; AAA51458.1; -.
DR    PIR; JC4023; JC4023.
DR    HSSP; P01137; 1KLA.
DR    InterPro; IPR002400; GF_cysknot.
DR    InterPro; IPR003911; TGF_TGFb.
DR    InterPro; IPR001839; TGFb.
DR    InterPro; IPR001111; TGFb_N.
DR    Pfam; PF00019; TGF-beta; 1.
DR    Pfam; PF00688; TGFb_propeptide; 1.
DR    PRINTS; PR00438; GFCYSKNOT.
DR    PRINTS; PR01423; TGFBETA.
DR    ProDom; PD000357; TGFb; 1.
DR    SMART; SM00204; TGFB; 1.
DR    PROSITE; PS00250; TGF_BETA_1; 1.
KW    Growth factor; Mitogen; Glycoprotein; Signal.
FT    SIGNAL          1      23      POTENTIAL.
FT    PROPEP          24     278     BY SIMILARITY.
FT    CHAIN           279     390     TRANSFORMING GROWTH FACTOR BETA 1.
FT    DISULFID        285     294     BY SIMILARITY.
FT    DISULFID        293     356     BY SIMILARITY.
FT    DISULFID        322     387     BY SIMILARITY.
FT    DISULFID        326     389     BY SIMILARITY.
FT    DISULFID        355     355     INTERCHAIN.
FT    CARBOHYD        82      82      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT    CARBOHYD       136     136     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT    CARBOHYD       176     176     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT    SITE            244     246     CELL ATTACHMENT SITE (POTENTIAL).
SQ    SEQUENCE       390 AA;  44185 MW;  EB4780E88B7B590E CRC64;

```

```

Query Match          93.4%;  Score 640;  DB 1;  Length 390;
Best Local Similarity 99.1%;  Pred. No. 1.5e-61;
Matches 112;  Conservative 1;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

```

```

Qy      68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     338 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 3

TGF1\_CERAE

```

ID   TGF1_CERAE          STANDARD;          PRT;          390 AA.
AC   P09533;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Transforming growth factor beta 1 precursor (TGF-beta 1).
GN   TGFb1.
OS   Cercopithecus aethiops (Green monkey) (Grivet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercopithecus.
OX   NCBI_TaxID=9534;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87246074; PubMed=3474130;
RA   Sharples K., Plowman G.D., Rose T.M., Twardzik D.R., Purchio A.F.;
RT   "Cloning and sequence analysis of simian transforming growth
RT   factor-beta cDNA.";
RL   DNA 6:239-244(1987).
CC   -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROL PROLIFERATION,
CC   DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY
CC   CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM HAVE
CC   SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC   ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A
CC   POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the TGF-beta family.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; M16658; AAA35369.1; -.
DR   PIR; A26960; A26960.
DR   HSSP; P01137; 1KLA.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR003911; TGF_TGFb.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 1.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   PRINTS; PR01423; TGFBETA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.

```

DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein; Signal.  
FT SIGNAL 1 16 POTENTIAL.  
FT PROPEP 17 278  
FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
FT DISULFID 285 294 BY SIMILARITY.  
FT DISULFID 293 356 BY SIMILARITY.  
FT DISULFID 322 387 BY SIMILARITY.  
FT DISULFID 326 389 BY SIMILARITY.  
FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
SQ SEQUENCE 390 AA; 44356 MW; DFF63E2BAB44320E CRC64;

Query Match 93.4%; Score 640; DB 1; Length 390;  
Best Local Similarity 99.1%; Pred. No. 1.5e-61;  
Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
:|||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
|||||||||||||||||||||||||||||||||||||||||||  
Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

#### RESULT 4

##### TGF1\_HUMAN

ID TGF1\_HUMAN STANDARD; PRT; 390 AA.  
AC P01137; Q9UCG4;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
GN TGFB1 OR TGFB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87174845; PubMed=3470709;  
RA Derynck R., Rhee L., Chen E.Y., van Tilburg A.;  
RT "Intron-exon structure of the human transforming growth factor-beta  
RT precursor gene.";  
RL Nucleic Acids Res. 15:3188-3189(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85296301; PubMed=3861940;  
RA Derynck R., Jarrett J.A., Chen E.Y., Eaton D.H., Bell J.R.,  
RA Assoian R.K., Roberts A.B., Sporn M.B., Goeddel D.V.;  
RT "Human transforming growth factor-beta complementary DNA sequence and  
RT expression in normal and transformed cells.";  
RL Nature 316:701-705(1985).

RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Duodenum, and Eye;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 279-390 FROM N.A.  
 RC TISSUE=Carcinoma;  
 RA Urushizaki Y., Niitsu Y., Terui T., Koshida Y., Mahara K., Kohgo Y.,  
 RA Urushizaki I., Takahashi Y., Ito H.;  
 RT "Cloning and expression of the gene for human transforming growth  
 RT factor-beta in Escherichia coli.";  
 RL Tumor Res. 22:41-55(1987).  
 RN [5]  
 RP SEQUENCE OF 279-329.  
 RC TISSUE=Bladder carcinoma;  
 RX MEDLINE=93229900; PubMed=8471846;  
 RA Bourdrel L., Lin C.-H., Lauren S.L., Elmore R.H., Sugarman B.J.,  
 RA Hu S., Westcott K.R.;  
 RT "Recombinant human transforming growth factor-beta 1: expression by  
 RT Chinese hamster ovary cells, isolation, and characterization.";  
 RL Protein Expr. Purif. 4:130-140(1993).  
 RN [6]  
 RP SEQUENCE OF 279-301.  
 RX MEDLINE=85131019; PubMed=2982829;  
 RA Massague J., Like B.;  
 RT "Cellular receptors for type beta transforming growth factor. Ligand  
 RT binding and affinity labeling in human and rodent cell lines.";  
 RL J. Biol. Chem. 260:2636-2645(1985).  
 RN [7]  
 RP STRUCTURE BY NMR OF 279-390.  
 RX MEDLINE=93144319; PubMed=8424942;  
 RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,  
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,  
 RA Torchia D.A.;  
 RT "Transforming growth factor beta 1: NMR signal assignments of the  
 RT recombinant protein expressed and isotopically enriched using Chinese  
 RT hamster ovary cells.";

RL Biochemistry 32:1152-1163(1993).  
 RN [8]  
 RP STRUCTURE BY NMR OF 279-390.  
 RX MEDLINE=93144320; PubMed=8424943;  
 RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,  
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,  
 RA Torchia D.A.;  
 RT "Transforming growth factor beta 1: secondary structure as determined  
 RT by heteronuclear magnetic resonance spectroscopy."  
 RL Biochemistry 32:1164-1171(1993).  
 RN [9]  
 RP STRUCTURE BY NMR OF 279-390.  
 RX MEDLINE=96266150; PubMed=8679613;  
 RA Hinck A.P., Archer S.J., Qian S.W., Roberts A.B., Sporn M.B.,  
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,  
 RA Torchia D.A.;  
 RT "Transforming growth factor beta 1: three-dimensional structure in  
 RT solution and comparison with the X-ray structure of transforming  
 RT growth factor beta 2."  
 RL Biochemistry 35:8517-8534(1996).  
 CC -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROLS PROLIFERATION,  
 CC DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY  
 CC CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF THEM HAVE  
 CC SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1 REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A  
 CC POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X05839; CAA29283.1; -.  
 DR EMBL; X05840; CAA29283.1; JOINED.  
 DR EMBL; X05843; CAA29283.1; JOINED.  
 DR EMBL; X05844; CAA29283.1; JOINED.  
 DR EMBL; X05849; CAA29283.1; JOINED.  
 DR EMBL; X05850; CAA29283.1; JOINED.  
 DR EMBL; X02812; CAA26580.1; ALT\_SEQ.  
 DR EMBL; BC001180; AAH01180.1; -.  
 DR EMBL; BC000125; AAH00125.1; -.  
 DR EMBL; BC022242; AAH22242.1; -.  
 DR EMBL; M38449; AAA36735.1; -.  
 DR PIR; A27513; WFHU2.  
 DR PDB; 1KLA; 17-AUG-96.  
 DR PDB; 1KLC; 17-AUG-96.  
 DR PDB; 1KLD; 17-AUG-96.  
 DR Genew; HGNC:11766; TGFB1.  
 DR MIM; 190180; -.  
 DR GO; GO:0006916; P:anti-apoptosis; TAS.  
 DR InterPro; IPR002400; GF\_cysknot.

DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal; 3D-structure.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 278  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294  
 FT DISULFID 293 356  
 FT DISULFID 322 387  
 FT DISULFID 326 389  
 FT DISULFID 355 355 INTERCHAIN.  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
 FT CONFLICT 10 10 L -> P (IN REF. 2).  
 FT CONFLICT 159 159 R -> RR (IN REF. 2).  
 FT STRAND 281 281  
 FT TURN 282 287  
 FT STRAND 294 296  
 FT STRAND 300 300  
 FT TURN 302 305  
 FT STRAND 313 313  
 FT STRAND 317 317  
 FT STRAND 321 323  
 FT HELIX 335 346  
 FT TURN 348 349  
 FT STRAND 358 370  
 FT TURN 371 372  
 FT STRAND 373 387  
 SQ SEQUENCE 390 AA; 44341 MW; 75391614250288FE CRC64;

Query Match 93.4%; Score 640; DB 1; Length 390;  
 Best Local Similarity 99.1%; Pred. No. 1.5e-61;  
 Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
 : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 5  
 TGF1\_PIG  
 ID TGF1\_PIG STANDARD; PRT; 390 AA.  
 AC P07200; P08832;

DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
 GN TGFB1.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RX MEDLINE=87174844; PubMed=3470708;  
 RA Derynck R., Rhee L.;  
 RT "Sequence of the porcine transforming growth factor-beta precursor.";  
 RL Nucleic Acids Res. 15:3187-3187(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT VAL-114.  
 RC STRAIN=Miniature swine;  
 RX MEDLINE=89054010; PubMed=2461367;  
 RA Kondaiah P., van Obberghen-Schilling E., Ludwig R.L., Dhar R.,  
 RA Sporn M.B., Roberts A.B.;  
 RT "cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.  
 RT Evidence for alternate splicing and polyadenylation.";  
 RL J. Biol. Chem. 263:18313-18317(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANT VAL-114.  
 RX MEDLINE=88335639; PubMed=3166520;  
 RA Jakowlew S.B., Dillard P.J., Sporn M.B., Roberts A.B.;  
 RT "Nucleotide sequence of chicken transforming growth factor-beta 1  
 RT (TGF-beta 1).";  
 RL Nucleic Acids Res. 16:8730-8730(1988).  
 RN [4]  
 RP SHOWS THAT REF.3 SEQUENCE IS FROM PIG.  
 RA Jakowlew S.B.;  
 RL Unpublished observations (MAR-1996).  
 RN [5]  
 RP SEQUENCE FROM N.A., AND VARIANT VAL-114.  
 RA Wimmers K., Chomdej S., Ponsuksili S., Schellander K.;  
 RT "Polymorphism in the porcine transforming growth factor beta 1  
 RT gene.";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP SEQUENCE OF 279-322.  
 RX MEDLINE=87102890; PubMed=2879635;  
 RA Cheifetz S., Weatherbee J.A., Tsang M.L.S., Anderson J.K., Mole J.E.,  
 RA Lucas R., Massague J.;  
 RT "The transforming growth factor-beta system, a complex pattern of  
 RT cross-reactive ligands and receptors.";  
 RL Cell 48:409-415(1987).  
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.



```

CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the TGF-beta family.
CC      -!- CAUTION: REF.3 SEQUENCE WHICH WAS SAID TO ORIGINATE FROM CHICKEN
CC          WHITE LEGHORN, SEEMS (REF.4) TO ORIGINATE FROM PIG.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Y00111; CAA68291.1; -.
DR      EMBL; M23703; AAA64616.1; -.
DR      EMBL; X12373; CAA30933.1; -.
DR      EMBL; AF461808; AAL57902.1; -.
DR      PIR; A27512; A27512.
DR      HSSP; P01137; 1KLA.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR003911; TGF_TGFb.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFbBETA.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFb; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
KW      Growth factor; Mitogen; Glycoprotein; Signal; Polymorphism.
FT      SIGNAL          1      23      POTENTIAL.
FT      PROPEP           24     278
FT      CHAIN           279     390      TRANSFORMING GROWTH FACTOR BETA 1.
FT      DISULFID        285     294      BY SIMILARITY.
FT      DISULFID        293     356      BY SIMILARITY.
FT      DISULFID        322     387      BY SIMILARITY.
FT      DISULFID        326     389      BY SIMILARITY.
FT      DISULFID        355     355      INTERCHAIN (BY SIMILARITY).
FT      CARBOHYD         82      82      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        136     136      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        176     176      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      SITE            244     246      CELL ATTACHMENT SITE (POTENTIAL).
FT      VARIANT         114     114      L -> V.
FT      CONFLICT         6       7      LR -> PG (IN REF. 3).
FT      CONFLICT        180     180      R -> G (IN REF. 3).
FT      CONFLICT        237     237      N -> NA (IN REF. 3).
SQ      SEQUENCE       390 AA;  44294 MW;  A6E2C3659FC384E6 CRC64;

```

```

Query Match          93.4%;  Score 640;  DB 1;  Length 390;
Best Local Similarity 99.1%;  Pred. No. 1.5e-61;
Matches 112;  Conservative 1;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67
      : |||||
Db      278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

```

```

Qy      68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      338 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 6

TGF1\_SHEEP

```

ID      TGF1_SHEEP      STANDARD;      PRT;      390 AA.
AC      P50414;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 1 precursor (TGF-beta 1).
GN      TGFB1.
OS      Ovis aries (Sheep).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC      Bovidae; Caprinae; Ovis.
OX      NCBI_TaxID=9940;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=95121932; PubMed=7821809;
RA      Woodall C.J., McLaren L.J., Watt N.J.;
RT      "Sequence and chromosomal localisation of the gene encoding ovine
RT      latent transforming growth factor-beta 1.";
RL      Gene 150:371-373(1994).
RN      [2]
RP      SEQUENCE OF 281-390 FROM N.A.
RC      STRAIN=Merino; TISSUE=Skin;
RX      MEDLINE=95268698; PubMed=7749621;
RA      Sutton R., Ward W.G., Raphael K.A., Cam G.R.;
RT      "Growth factor expression in skin during wool follicle development.";
RL      Comp. Biochem. Physiol. 110B:697-705(1995).
CC      -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC      PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC      TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC      HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC      A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC      -!- SUBUNIT: Homodimer; disulfide-linked.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the TGF-beta family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X76916; CAA54242.1; -.
DR      EMBL; L36038; AAA31526.1; -.
DR      PIR; I46463; I46463.
DR      HSSP; P01137; 1KLA.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR003911; TGF_TGFB.

```

DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 278 POTENTIAL.  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294 BY SIMILARITY.  
 FT DISULFID 293 356 BY SIMILARITY.  
 FT DISULFID 322 387 BY SIMILARITY.  
 FT DISULFID 326 389 BY SIMILARITY.  
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 390 AA; 44291 MW; 1C247299484D0E57 CRC64;

Query Match 93.4%; Score 640; DB 1; Length 390;  
 Best Local Similarity 99.1%; Pred. No. 1.5e-61;  
 Matches 112; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 338 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

# RESULT 7

TGF1\_MOUSE  
 ID TGF1\_MOUSE STANDARD; PRT; 390 AA.  
 AC P04202;  
 DT 20-MAR-1987 (Rel. 04, Created)  
 DT 20-MAR-1987 (Rel. 04, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
 GN TGFb1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86168129; PubMed=3007454;  
 RA Derynck R., Jarrett J.A., Chen E.Y., Goeddel D.V.;  
 RT "The murine transforming growth factor-beta precursor."  
 RL J. Biol. Chem. 261:4377-4379(1986).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RX MEDLINE=96096545; PubMed=8522200;  
 RA Guron C., Sudarshan C., Raghov R.;  
 RT "Molecular organization of the gene encoding murine transforming  
 RT growth factor beta 1.";  
 RL Gene 165:325-326(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6, and NOD/LT; TISSUE=Spleen;  
 RA Poirot L., Benoist C., Mathis D.;  
 RT "Transforming growth factor-beta 1 sequence and expression: no  
 RT difference between NOD/Lt and C57Bl/6 mouse strains.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M13177; AAA40423.1; -.  
 DR EMBL; L42462; AAB00138.1; -.  
 DR EMBL; L42456; AAB00138.1; JOINED.  
 DR EMBL; L42457; AAB00138.1; JOINED.  
 DR EMBL; L42458; AAB00138.1; JOINED.  
 DR EMBL; L42459; AAB00138.1; JOINED.  
 DR EMBL; L42460; AAB00138.1; JOINED.  
 DR EMBL; L42461; AAB00138.1; JOINED.  
 DR EMBL; AJ009862; CAA08900.1; -.  
 DR PIR; A01396; WFMS2.  
 DR HSSP; P01137; 1KLA.  
 DR MGD; MGI:98725; Tgfb1.  
 DR GO; GO:0005578; C:extracellular matrix; IDA.  
 DR GO; GO:0006954; P:inflammatory response; IMP.  
 DR GO; GO:0007515; P:lymph gland development; IMP.  
 DR GO; GO:0008220; P:necrosis; IMP.  
 DR GO; GO:0016202; P:regulation of myogenesis; IDA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.

DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 278  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294 BY SIMILARITY.  
 FT DISULFID 293 356 BY SIMILARITY.  
 FT DISULFID 322 387 BY SIMILARITY.  
 FT DISULFID 326 389 BY SIMILARITY.  
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 390 AA; 44310 MW; 4381A51B711D689E CRC64;

Query Match 93.0%; Score 637; DB 1; Length 390;  
 Best Local Similarity 98.2%; Pred. No. 3.1e-61;  
 Matches 111; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :|||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
 Qy 68 KVLALYNQHNPASAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
 |||||||||||||||||||:|||||||||||||||||||||||||||  
 Db 338 KVLALYNQHNPASAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 8  
 TGF1\_RAT  
 ID TGF1\_RAT STANDARD; PRT; 390 AA.  
 AC P17246;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
 GN TGFB1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Heart;  
 RX MEDLINE=90272425; PubMed=2349108;  
 RA Qian S.W., Kondaiah P., Roberts A.B., Sporn M.B.;  
 RT "cDNA cloning by PCR of rat transforming growth factor beta-1.";  
 RL Nucleic Acids Res. 18:3059-3059(1990).  
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.

CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X52498; CAA36741.1; -.  
 DR PIR; S10219; S10219.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 278  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294 BY SIMILARITY.  
 FT DISULFID 293 356 BY SIMILARITY.  
 FT DISULFID 322 387 BY SIMILARITY.  
 FT DISULFID 326 389 BY SIMILARITY.  
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 390 AA; 44329 MW; 5E21108ED50D853C CRC64;

Query Match 93.0%; Score 637; DB 1; Length 390;  
 Best Local Similarity 98.2%; Pred. No. 3.1e-61;  
 Matches 111; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||:||||||||||||||||||||||||||||||||||  
 Db 338 KVLALYNQHNPASASPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 9  
 TGF1\_HORSE  
 ID TGF1\_HORSE STANDARD; PRT; 390 AA.

AC O19011;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
 GN TGFB1.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lymph node;  
 RX MEDLINE=98185507; PubMed=9524819;  
 RA Penha-Goncalves M.N., Onions D.E., Nicolson L.;  
 RT "Cloning and sequencing of equine transforming growth factor-beta 1  
 RT (TGF beta-1) cDNA.";  
 RL DNA Seq. 7:375-378(1997).  
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X99438; CAA67801.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 278 BY SIMILARITY.  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294 BY SIMILARITY.  
 FT DISULFID 293 356 BY SIMILARITY.  
 FT DISULFID 322 387 BY SIMILARITY.  
 FT DISULFID 326 389 BY SIMILARITY.

FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 390 AA; 43974 MW; A86D715F44549691 CRC64;

Query Match 92.8%; Score 636; DB 1; Length 390;  
 Best Local Similarity 98.2%; Pred. No. 4e-61;  
 Matches 111; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : ||||||||||||||||||||||||||||||||||||||||||||  
 Db 278 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 338 KVLALYNQHNPASAAAPCCVPQVLEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

# RESULT 10

## TGF1\_CAVPO

ID TGF1\_CAVPO STANDARD; PRT; 390 AA.  
 AC Q9Z1Y6; Q9QZB3; Q9R148;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
 GN TGFBI.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hartley;  
 RA Jeevan A., McMurray D.N., Yoshimura T.;  
 RT "Guinea pig transforming growth factor-beta in peritoneal exudates  
 RT after BCG vaccination."  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE OF 265-382 FROM N.A.  
 RX MEDLINE=99144670; PubMed=10025978;  
 RA Scarozza A.M., Ramsingh A.I., Wicher V., Wicher K.;  
 RT "Spontaneous cytokine gene expression in normal guinea pig blood and  
 RT tissues."  
 RL Cytokine 10:851-859(1998).  
 RN [3]  
 RP SEQUENCE OF 279-371 FROM N.A.  
 RC STRAIN=Hartley; TISSUE=Trachea;  
 RA Morishima Y., Uchida Y., Nomura A., Ishii Y., Sakamoto T.,  
 RA Sekizawa K.;  
 RT "Guinea-pig transforming growth factor-beta expression in injured  
 RT tracheal epithelium."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROLS PROLIFERATION,  
 CC DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY  
 CC CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF THEM HAVE



CC        SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1 REGULATES THE  
CC        ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A  
CC        POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
CC    -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).  
CC    -!- SUBCELLULAR LOCATION: Secreted.  
CC    -!- SIMILARITY: Belongs to the TGF-beta family.  
CC    -----  
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC    the European Bioinformatics Institute. There are no restrictions on its  
CC    use by non-profit institutions as long as its content is in no way  
CC    modified and this statement is not removed. Usage by and for commercial  
CC    entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC    or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC    -----  
DR    EMBL; AF191297; AAF02780.1; -.  
DR    EMBL; AF097509; AAC83807.1; -.  
DR    EMBL; AF169347; AAD49347.1; -.  
DR    HSSP; P01137; 1KLA.  
DR    InterPro; IPR002400; GF\_cysknot.  
DR    InterPro; IPR003911; TGF\_TGFB.  
DR    InterPro; IPR001839; TGFB.  
DR    InterPro; IPR001111; TGFB\_N.  
DR    Pfam; PF00019; TGF-beta; 1.  
DR    Pfam; PF00688; TGFB\_propeptide; 1.  
DR    PRINTS; PR00438; GFCYSKNOT.  
DR    PRINTS; PR01423; TGFBETA.  
DR    ProDom; PD000357; TGFB; 1.  
DR    SMART; SM00204; TGFB; 1.  
DR    PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW    Growth factor; Mitogen; Glycoprotein; Signal.  
FT    SIGNAL            1        24        POTENTIAL.  
FT    PROPEP            25       278       POTENTIAL.  
FT    CHAIN            279       390       TRANSFORMING GROWTH FACTOR BETA 1.  
FT    DISULFID        285       294       BY SIMILARITY.  
FT    DISULFID        293       356       BY SIMILARITY.  
FT    DISULFID        322       387       BY SIMILARITY.  
FT    DISULFID        326       389       BY SIMILARITY.  
FT    DISULFID        355       355       INTERCHAIN (BY SIMILARITY).  
FT    CARBOHYD        82        82       N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT    CARBOHYD       136       136       N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT    CARBOHYD       176       176       N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT    SITE            244       246       CELL ATTACHMENT SITE (POTENTIAL).  
FT    CONFLICT        279       279       G -> P (IN REF. 3).  
FT    CONFLICT        286       286       F -> S (IN REF. 2).  
FT    CONFLICT        309       309       K -> E (IN REF. 2).  
FT    CONFLICT        322       322       C -> R (IN REF. 2).  
FT    CONFLICT        350       350       A -> G (IN REF. 2).  
SQ    SEQUENCE       390 AA;   44328 MW;   1539F849BA0C0FF1 CRC64;

Query Match                    91.7%;   Score 628;   DB 1;   Length 390;  
Best Local Similarity       97.3%;   Pred. No. 2.9e-60;  
Matches 110;   Conservative    1;   Mismatches    2;   Indels       0;   Gaps       0;

Qy                    8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
: ||||||||||||||||||||||||||||||||||||||||||||  
Db                    278 RGLDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 337

Qy 68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 338 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKAKVEQLSNMIVRSCKCS 390

RESULT 11

TGF1\_CHICK

ID TGF1\_CHICK STANDARD; PRT; 373 AA.  
 AC P09531;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 1 precursor (TGF-beta 1) (TGF-beta 4)  
 DE (Fragment).  
 GN TGFB1.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=White leghorn;  
 RX MEDLINE=89112198; PubMed=2464131;  
 RA Jakowlew S.B., Dillard P.J., Sporn M.B., Roberts A.B.;  
 RT "Complementary deoxyribonucleic acid cloning of a messenger  
 RT ribonucleic acid encoding transforming growth factor beta 4 from  
 RT chicken embryo chondrocytes.";  
 RL Mol. Endocrinol. 2:1186-1195(1988).  
 RN [2]  
 RP REVISIONS.  
 RX MEDLINE=92357039; PubMed=1353860;  
 RA Burt D.W., Jakowlew S.B.;  
 RT "Correction: a new interpretation of a chicken transforming growth  
 RT factor-beta 4 complementary DNA.";  
 RL Mol. Endocrinol. 6:989-992(1992).  
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M31160; AAB05637.1; -.  
 DR PIR; A41918; A41918.

DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 1 POTENTIAL.  
 FT PROPEP 2 259 POTENTIAL.  
 FT CHAIN 260 373 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 266 277 BY SIMILARITY.  
 FT DISULFID 276 339 BY SIMILARITY.  
 FT DISULFID 305 370 BY SIMILARITY.  
 FT DISULFID 309 372 BY SIMILARITY.  
 FT DISULFID 338 338 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 224 226 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 373 AA; 42634 MW; 9903F3479C8552E5 CRC64;

Query Match 79.1%; Score 542; DB 1; Length 373;  
 Best Local Similarity 82.3%; Pred. No. 5.2e-51;  
 Matches 93; Conservative 9; Mismatches 9; Indels 2; Gaps 1;

Qy 10 LDTNYCF--SSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 |||:| | : ||||| ||||| ||||| ||||| ||||| ||||| :  
 Db 261 LDTDYCFGPGTDEKNCCVRPLYIDFRKDLQWKWIHEPKGYMANFCMGPCPYIWSADTQYT 320  
  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :|||:| | :|||:| | :|||  
 Db 321 KVLALYNQHNPASAAAPCCVPQTLDPPIIYVGRNVRVEQLSNMVVRACKCS 373

# RESULT 12

## TGF1\_XENLA

ID TGF1\_XENLA STANDARD; PRT; 382 AA.  
 AC P16176;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 1 precursor (TGF-beta 1) (TGF-beta 5).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90110090; PubMed=2295601;  
 RA Kondaiah P., Sands M.J., Smith J.M., Fields A., Roberts A.B.,  
 RA Sporn M.B., Melton D.A.;

RT "Identification of a novel transforming growth factor-beta (TGF-beta  
RT 5) mRNA in *Xenopus laevis*.";  
RL J. Biol. Chem. 265:1089-1093(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Vempati U.D., Kondaiah P.;  
RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: IMPORTANT ROLE IN CERTAIN ASPECTS OF DIFFERENTIATION.  
CC -!- SUBUNIT: Homodimer; disulfide-linked.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the TGF-beta family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J05180; AAA49968.1; -.  
DR EMBL; AF009335; AAB64441.1; -.  
DR EMBL; AF009331; AAB64441.1; JOINED.  
DR EMBL; AF009332; AAB64441.1; JOINED.  
DR EMBL; AF009333; AAB64441.1; JOINED.  
DR EMBL; AF009334; AAB64441.1; JOINED.  
DR PIR; A34929; B61036.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR003911; TGF\_TGFB.  
DR InterPro; IPR001839; TGFB.  
DR InterPro; IPR001111; TGFB\_N.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFB\_propeptide; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRINTS; PR01423; TGFBETA.  
DR ProDom; PD000357; TGFB; 1.  
DR SMART; SM00204; TGFB; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein; Signal.  
FT SIGNAL 1 21 POTENTIAL.  
FT PROPEP 22 270  
FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA 1.  
FT DISULFID 277 286 BY SIMILARITY.  
FT DISULFID 285 348 BY SIMILARITY.  
FT DISULFID 314 379 BY SIMILARITY.  
FT DISULFID 318 381 BY SIMILARITY.  
FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT SITE 234 236 CELL ATTACHMENT SITE (POTENTIAL).  
SQ SEQUENCE 382 AA; 44200 MW; 1034621C917AAE15 CRC64;

Query Match 73.7%; Score 505; DB 1; Length 382;  
Best Local Similarity 74.3%; Pred. No. 5.1e-47;  
Matches 84; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : : ||| : ||||: |||:||||||| ||:||| |||||:|||||  
 Db 270 RGVGQEYCFGNNGPNCCVKPLYINFRKDLGWKWIHEPKGYEANYCLGNCPYIWSMDTQYS 329

Qy 68 KVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 |||:||||:||||| :||||| |||||:||||| |||||:||||| ||  
 Db 330 KVLSLYNQNNPGASISPCCVDPVLEPLPIIYYVGRTAKVEQLSNMVVRSCNCS 382

# RESULT 13

## TGF3\_MOUSE

ID TGF3\_MOUSE STANDARD; PRT; 410 AA.  
 AC P17125;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 3 precursor (TGF-beta 3).  
 GN TGFB3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90190650; PubMed=2628730;  
 RA Miller D.A., Lee A., Matsui Y., Chen E.Y., Moses H.L., Derynck R.;  
 RT "Complementary DNA cloning of the murine transforming growth  
 RT factor-beta 3 (TGF beta 3) precursor and the comparative expression  
 RT of TGF beta 3 and TGF beta 1 messenger RNA in murine embryos and  
 RT adult tissues.";  
 RL Mol. Endocrinol. 3:1926-1934(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91000714; PubMed=2206556;  
 RA Denhez F., Lafyatis R., Kondaiah P., Roberts A.B., Sporn M.B.;  
 RT "Cloning by polymerase chain reaction of a new mouse TGF-beta,  
 RT mTGF-beta 3.";  
 RL Growth Factors 3:139-146(1990).  
 CC -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M32745; AAA40422.1; -.  
 DR PIR; A41397; A41397.  
 DR HSSP; P10600; 1TGJ.  
 DR MGD; MGI:98727; Tgfb3.  
 DR InterPro; IPR002400; GF\_cysknot.

DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 298 POTENTIAL.  
 FT CHAIN 299 410 TRANSFORMING GROWTH FACTOR BETA 3.  
 FT DISULFID 305 314 BY SIMILARITY.  
 FT DISULFID 313 376 BY SIMILARITY.  
 FT DISULFID 342 407 BY SIMILARITY.  
 FT DISULFID 346 409 BY SIMILARITY.  
 FT DISULFID 375 375 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 259 261 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 410 AA; 46884 MW; 250F7048CA432BD6 CRC64;

Query Match 73.3%; Score 502; DB 1; Length 410;  
 Best Local Similarity 76.1%; Pred. No. 1.2e-46;  
 Matches 86; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||||||| : |:||||| |||||:|||||:|||||:||||| |||||: || :|  
 Db 298 RALDTNYCFRNLEENCCVRPLYIDFRQDLGWKQVHEPKGYANFCSGPCPYLRSADTTTHS 357  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
 || ||| || |||:||||| |||| |:||||| |||||:|:|||||  
 Db 358 TVLGLYNTLNPEASASPCCVPQDLEPLTILYVGRTPKVEQLSNMVVKSCCKCS 410

#### RESULT 14

##### TGF3\_HUMAN

ID TGF3\_HUMAN STANDARD; PRT; 412 AA.  
 AC P10600;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Transforming growth factor beta 3 precursor (TGF-beta 3).  
 GN TGFB3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88263019; PubMed=3164476;  
 RA ten Dijke P., Hansen P., Iwata K., Pieler C., Foulkes J.G.;  
 RT "Identification of another member of the transforming growth factor  
 RT type beta gene family.";

RL Proc. Natl. Acad. Sci. U.S.A. 85:4715-4719(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=89091120; PubMed=3208746;  
 RA Derynck R., Lindquist P.B., Lee A., Wen D., Tamm J., Graycar J.L.,  
 RA Rhee L., Mason A.J., Miller D.A., Coffey R.J., Moses H.L., Chen E.Y.;  
 RT "A new type of transforming growth factor-beta, TGF-beta 3."  
 RL EMBO J. 7:3737-3743(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Madan A., Rowen L., Qin S., Dickhoff R., Shaffer T., James R.,  
 RA Abbasi N., Loretz C., Madan A., Dors M., Dahl T., Hall J., Lasky S.,  
 RA Hood L.;  
 RT "Complete genomic sequence of human transforming growth factor-beta  
 RT 3."  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,  
 RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,  
 RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 301-412.  
 RX MEDLINE=96416253; PubMed=8819159;  
 RA Mittl P.R., Priestle J.P., Cox D.A., McMaster G., Cerletti N.,  
 RA Grutter M.G.;  
 RT "The crystal structure of TGF-beta 3 and comparison to TGF-beta 2:  
 RT implications for receptor binding."  
 RL Protein Sci. 5:1261-1271(1996).  
 CC -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J03241; AAA61161.1; -.  
 DR EMBL; X14149; CAA32362.1; -.  
 DR EMBL; X14885; CAA33024.1; ALT\_INIT.  
 DR EMBL; X14886; CAA33024.1; JOINED.  
 DR EMBL; X14887; CAA33024.1; JOINED.  
 DR EMBL; X14888; CAA33024.1; JOINED.  
 DR EMBL; X14889; CAA33024.1; JOINED.  
 DR EMBL; X14890; CAA33024.1; JOINED.  
 DR EMBL; X14891; CAA33024.1; JOINED.  
 DR EMBL; AF107885; AAC79727.1; -.  
 DR EMBL; AY140241; AAM96819.1; -.  
 DR PIR; A36169; A36169.  
 DR PDB; 1TGT; 11-JAN-97.

```

DR   PDB; 1TGK; 12-MAR-97.
DR   PDB; 1KTZ; 27-FEB-02.
DR   Genew; HGNC:11769; TGFB3.
DR   MIM; 190230; -.
DR   GO; GO:0005160; F:transforming growth factor-beta receptor li. . .; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR003911; TGF_TGFB.
DR   InterPro; IPR001839; TGFB.
DR   InterPro; IPR001111; TGFB_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFB_propeptide; 1.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   PRINTS; PR01423; TGFBETA.
DR   ProDom; PD000357; TGFB; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Growth factor; Mitogen; Glycoprotein; Signal; 3D-structure.
FT   SIGNAL          1      20      POTENTIAL.
FT   PROPEP          21     300
FT   CHAIN           301     412      TRANSFORMING GROWTH FACTOR BETA 3.
FT   DISULFID        307     316
FT   DISULFID        315     378
FT   DISULFID        344     409
FT   DISULFID        348     411
FT   DISULFID        377     377      INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD         74      74      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD        135     135      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD        142     142      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   SITE            261     263      CELL ATTACHMENT SITE (POTENTIAL).
FT   STRAND           303     303
FT   HELIX            304     309
FT   STRAND           316     318
FT   STRAND           321     323
FT   HELIX            324     328
FT   TURN             331     332
FT   STRAND           333     335
FT   STRAND           338     340
FT   STRAND           343     345
FT   TURN             350     353
FT   STRAND           354     354
FT   HELIX            357     368
FT   TURN             370     371
FT   STRAND           378     380
FT   STRAND           383     392
FT   TURN             393     394
FT   STRAND           395     406
FT   STRAND           408     412
SQ   SEQUENCE        412 AA;  47328 MW;  3CAD3548D3AEA178 CRC64;

Query Match          73.3%;  Score 502;  DB 1;  Length 412;
Best Local Similarity 76.1%;  Pred. No. 1.2e-46;
Matches 86;  Conservative 12;  Mismatches 15;  Indels 0;  Gaps 0;

Qy          8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67

```







Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result			%		Query		ID	Description
	No.	Score	Match	Length	DB			
1	638	93.1	112	6	002730		002730	oryctolagus
2	637	93.0	368	11	Q8R4D9		Q8r4d9	sigmodon hi
3	628	91.7	390	6	Q9TUM8		Q9tum8	equus cabal
4	625	91.2	130	11	Q08714		Q08714	mesocricetu
5	622	90.8	124	6	Q95N80		Q95n80	canis famil
6	554	80.9	101	11	Q9R184		Q9rl84	meriones un
7	502	73.3	362	11	Q99K17		Q99k17	mus musculu
8	502	73.3	412	11	Q91YU7		Q91yu7	mus musculu
9	482	70.4	255	11	Q921T1		Q921t1	mus musculu
10	482	70.4	414	11	Q91VP5		Q91vp5	mus musculu
11	471.5	68.8	200	13	Q90YF1		Q90yf1	pleuronecte
12	470	68.6	224	11	Q8CDZ9		Q8cdz9	mus musculu
13	470	68.6	382	13	O93449		O93449	oncorhynchu
14	465.5	68.0	382	13	Q9PWA9		Q9pwa9	morone chry
15	464.5	67.8	379	13	Q8JHF5		Q8jhf5	sparus aura
16	460.5	67.2	379	13	Q8AXK8		Q8axk8	sparus aura
17	440	64.2	361	13	Q98854		Q98854	cyprinus ca
18	436	63.6	399	11	Q9ERB7		Q9erb7	mesocricetu
19	429	62.6	376	13	Q9PTQ2		Q9ptq2	cyprinus ca
20	413	60.3	88	13	Q90YF5		Q90yf5	pleuronecte

21	402	58.7	88	13	Q90YF7	Q90yf7 oncorhynchu
22	397	58.0	88	13	Q90ZE7	Q90ze7 acipenser b
23	393	57.4	87	13	O42306	O42306 carassius a
24	383	55.9	91	6	Q9MYZ1	Q9myz1 capra hircu
25	373	54.5	86	6	Q28241	Q28241 cervus elap
26	369	53.9	179	13	Q90YF2	Q90yf2 pleuronecte
27	358	52.3	81	6	Q9N1S3	Q9nls3 capreolus c
28	317	46.3	77	13	Q90YF8	Q90yf8 oncorhynchu
29	302	44.1	50	6	Q28240	Q28240 cervus elap
30	301	43.9	62	13	Q90ZJ7	Q90zj7 anguilla an
31	300	43.8	62	13	Q90YF4	Q90yf4 pleuronecte
32	287	41.9	62	13	Q9DEP5	Q9dep5 scophthalmu
33	283	41.3	62	13	Q90ZJ8	Q90zj8 anguilla an
34	283	41.3	62	13	Q90YF9	Q90yf9 oncorhynchu
35	273	39.9	62	13	Q90YF3	Q90yf3 pleuronecte
36	202.5	29.6	374	13	Q8JFS0	Q8jfs0 brachydanio
37	202.5	29.6	385	13	Q90W05	Q90w05 sparus aura
38	202.5	29.6	389	13	Q90YY0	Q90yy0 ictalurus p
39	200	29.2	426	4	Q9HBP0	Q9hbp0 homo sapien
40	199.5	29.1	373	13	Q90ZD2	Q90zd2 oncorhynchu
41	199.5	29.1	373	13	Q90ZD1	Q90zd1 oncorhynchu
42	199.5	29.1	373	13	Q9DDI8	Q9ddi8 salmo salar
43	199.5	29.1	376	13	Q98TB4	Q98tb4 oreochromis
44	199.5	29.1	376	13	Q90WC9	Q90wc9 morone saxa
45	199.5	29.1	376	13	Q90WC8	Q90wc8 morone amer

# ALIGNMENTS

## RESULT 1

O02730

ID O02730 PRELIMINARY; PRT; 112 AA.

AC O02730; O97501;

DT 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).

GN TGFBI OR TGF-BETA-1.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI\_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RA Taylor T.K., James E.R., McGonigle S., Yoho E.R.;

RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.

RN [2]

RP SEQUENCE OF 2-99 FROM N.A.

RA Inoue K., Kawabe Y., Kodama T.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF  
 CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1  
 CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND  
 CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.

CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AF000133; AAB53806.1; -.  
 DR EMBL; AB020217; BAA36950.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein.  
 FT NON\_TER 1 1  
 FT CHAIN 1 112 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 7 16 BY SIMILARITY.  
 FT DISULFID 15 78 BY SIMILARITY.  
 FT DISULFID 44 109 BY SIMILARITY.  
 FT DISULFID 48 111 BY SIMILARITY.  
 FT DISULFID 77 77 INTERCHAIN (BY SIMILARITY).  
 FT CONFLICT 2 3 LD -> FS (IN REF. 2).  
 FT CONFLICT 85 92 PLPIVYYV -> ATAHVRTTL (IN REF. 2).  
 SQ SEQUENCE 112 AA; 12795 MW; 53C5B7D46355A6F3 CRC64;

Query Match 93.1%; Score 638; DB 6; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-66;  
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK 68  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK 60  
 Qy 69 VLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 VLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112

## RESULT 2

Q8R4D9

ID Q8R4D9 PRELIMINARY; PRT; 368 AA.  
 AC Q8R4D9;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta-1 protein (Fragment).  
 GN TGFB1.  
 OS Sigmodon hispidus (Hispid cotton rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
 OC Sigmodon.  
 OX NCBI\_TaxID=42415;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Blanco J.C., Pletneva L.M., Prince G.A.;  
 RT "Cotton rat cytokines, chemokines, and interferons."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

DR EMBL; AF480858; AAL87199.1; -.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 368 AA; 41905 MW; A5C91207B0468B4A CRC64;

Query Match 93.0%; Score 637; DB 11; Length 368;  
 Best Local Similarity 98.2%; Pred. No. 1.4e-65;  
 Matches 111; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 256 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 315  
 QY 68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||:||||||||||||||||||||||||||||||||||  
 Db 316 KVLALYNQHNP GASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 368

# RESULT 3

Q9TUM8

ID Q9TUM8 PRELIMINARY; PRT; 390 AA.  
 AC Q9TUM8;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta 1.  
 GN TGFb1.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nixon A.J., Brower-Toland B.T., Sandell L.J.;  
 RT "Molecular cloning of equine transforming growth factor beta 1 reveals  
 RT a divergent nucleotide structure that encodes a novel bioactive  
 RT peptide among mammalian species."  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AF175709; AAD49431.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFbBETA.



CC -!- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.  
 DR EMBL; X60296; CAA42838.1; -.  
 DR EMBL; AF046214; AAC40099.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR001839; TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein.  
 FT NON\_TER 1 1  
 FT PROPEP <1 18  
 FT CHAIN 19 130 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 25 34 BY SIMILARITY.  
 FT DISULFID 33 96 BY SIMILARITY.  
 FT DISULFID 66 129 BY SIMILARITY.  
 FT DISULFID 95 95 INTERCHAIN (BY SIMILARITY).  
 FT CONFLICT 93 93 G -> S (IN REF. 2).  
 SQ SEQUENCE 130 AA; 14997 MW; 8B41DD6CF39CCA77 CRC64;

Query Match 91.2%; Score 625; DB 11; Length 130;  
 Best Local Similarity 97.3%; Pred. No. 1.1e-64;  
 Matches 110; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 18 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 77  
 Qy 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 78 KVLALYNQHNPASAGPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSYKCS 130

# RESULT 5

Q95N80

ID Q95N80 PRELIMINARY; PRT; 124 AA.  
 AC Q95N80;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta 1 (Fragment).  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fonfara S., Groene A., Baumgaertner W.;  
 RT "Sequence of canine transforming growth factor beta 1 mRNA in DH82-  
 RT cells.";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AF349538; AAK54072.1; -.  
 DR InterPro; IPR001839; TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.



DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 124 124  
 SQ SEQUENCE 124 AA; 14329 MW; 21D185218E5556DB CRC64;

Query Match 90.8%; Score 622; DB 6; Length 124;  
 Best Local Similarity 99.1%; Pred. No. 2.3e-64;  
 Matches 109; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 : ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 15 RALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 74

Qy 68 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSC 117  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 75 KVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSC 124

# RESULT 6

Q9R184

ID Q9R184 PRELIMINARY; PRT; 101 AA.  
 AC Q9R184;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).  
 GN TGFB1 OR TGF-BETA.  
 OS Meriones unguiculatus (Mongolian jird).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
 OC Meriones.  
 OX NCBI\_TaxID=10047;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LYMPH NODE;  
 RA Rao U.R., Klei T.R.;  
 RT "cDNA cloning of gerbil transforming growth factor-beta by PCR."  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF  
 CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1  
 CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND  
 CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AF161218; AAD45726.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR001839; TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein.  
 FT NON\_TER 1 1  
 FT CHAIN <1 >101 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 1 10 BY SIMILARITY.

FT	DISULFID	9	72	BY SIMILARITY.
FT	DISULFID	71	71	INTERCHAIN (BY SIMILARITY).
FT	NON_TER	101	101	
SQ	SEQUENCE	101 AA;	11724 MW;	ABF1CFDA264AEFED CRC64;

Qy 15 CFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYN 74  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1 CFSSTEKNCCVRQLYRDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLHTQYSKVLALYN 60

## RESULT 7

```

ID      Q99K17          PRELIMINARY;          PRT;      362 AA.
AC      Q99K17;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Similar to transforming growth factor, beta 3 (Fragment).
GN      TGFB3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Strausberg R.;
RL      Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC      -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR      EMBL; BC005513; AAH05513.1; -.
DR      HSSP; P10600; 1TGJ.
DR      MGD; MGI:98727; Tgfb3.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      InterPro; IPR003911; TGF_TGFb.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFBETA.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFB; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
FT      NON_TER      1      1
SQ      SEQUENCE      362 AA;  41486 MW;  0808E46180FDAE70 CRC64;

```

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||||||| : |:||||| |||||:|||||:|||||:|||||: ||| :|  
 Db 250 RALDTNYCFRNLEENCCVRPLYIDFRQDLGWKWWHEPKGYANFCSGPCPYLRSADTTHS 309

Qy 68 KVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 || ||| || |||:||||| |||| |:||||| |||||:|:|||||  
 Db 310 TVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKCS 362

RESULT 8

Q91YU7

ID Q91YU7 PRELIMINARY; PRT; 412 AA.  
 AC Q91YU7;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor, beta 3.  
 GN TGFB3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; BC014690; AAH14690.1; -.  
 DR MGD; MGI:98727; Tgfb3.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGfb.  
 DR InterPro; IPR001111; TGfb\_N.  
 DR InterPro; IPR003911; TGF\_TGfb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGfb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGfb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 SQ SEQUENCE 412 AA; 47144 MW; F3EB65D046DF32AD CRC64;

Query Match 73.3%; Score 502; DB 11; Length 412;  
 Best Local Similarity 76.1%; Pred. No. 7.7e-50;  
 Matches 86; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

Qy 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||||||| : |:||||| |||||:|||||:|||||:|||||: ||| :|  
 Db 300 RALDTNYCFRNLEENCCVRPLYIDFRQDLGWKWWHEPKGYANFCSGPCPYLRSADTTHS 359

Qy 68 KVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 || ||| || |||:||||| |||| |:||||| |||||:|:|||||  
 Db 360 TVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKCS 412

RESULT 9

Q921T1



RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast tumor;  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; BC011170; AAH11170.1; -.  
 DR MGD; MGI:98726; Tgfb2.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 SQ SEQUENCE 414 AA; 47588 MW; DB37A7C38881F286 CRC64;

Query Match 70.4%; Score 482; DB 11; Length 414;  
 Best Local Similarity 70.8%; Pred. No. 1.6e-47;  
 Matches 80; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||| ||| : : |||:| |||||:| ||||| |||||:| ||| | |||:| |||:|  
 Db 302 RALDAAAYCFRNVQDNCCLRPLYIDFKRDLGWKWIHEPKGYNANFCAGACPYLWSSDTQHT 361  
 QY 68 KVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
 |||:| || |||:| ||| | ||| |:| |:| ||:| |||||:| |||||  
 Db 362 KVLSLYNTINPEASASPCCVSQDLEPLTILYYIGNTPKIEQLSNMIVKSKCS 414

# RESULT 11

Q90YF1

ID Q90YF1 PRELIMINARY; PRT; 200 AA.  
 AC Q90YF1;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta 1 (Fragment).  
 GN TGF-BETA1.  
 OS Pleuronectes platessa (Plaice).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
 OC Pleuronectoidei; Pleuronectidae; Pleuronectes.  
 OX NCBI\_TaxID=8262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laing K.J.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20394651; PubMed=10938738;  
 RA Laing K.J., Cunningham C., Secombes C.J.;

RT "Genes for three different isoforms of transforming growth factor-beta  
RT are present in plaice (*Pleuronectes platessa*) DNA.";  
RL Fish and Shellfish Immunol. 10:261-271(2000).  
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
DR EMBL; AJ318933; CAC60268.1; -.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR001839; TGFb.  
DR Pfam; PF00019; TGF-beta; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR ProDom; PD000357; TGFb; 1.  
DR SMART; SM00204; TGFb; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
FT NON\_TER 1 1  
FT NON\_TER 200 200  
SQ SEQUENCE 200 AA; 22851 MW; 4876FEB6A263B4CC CRC64;

Query Match 68.8%; Score 471.5; DB 13; Length 200;  
Best Local Similarity 73.4%; Pred. No. 1.2e-46;  
Matches 80; Conservative 14; Mismatches 14; Indels 1; Gaps 1;

Qy 12 TNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVL 71  
| : | : || ||||:||||||| |||||:| | |||: : :||:|  
Db 93 TDTCTAQTE-TCCVRKLYIDFRKDLGWKWIHKPTGYHANYCMGSCTYIWNNAENKYSQILA 151  
Qy 72 LYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120  
|| ||||| |||||:||||: ||||| |:||||  
Db 152 LYKHHNPGASAPCCVPQALEPLPIVYVGRQHKVEQLSNMSVKSCCKCS 200

# RESULT 12

Q8CDZ9

ID Q8CDZ9 PRELIMINARY; PRT; 224 AA.  
AC Q8CDZ9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Transforming growth factor (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
DR EMBL; AK029306; BAC26384.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 224 AA; 25538 MW; 9264D6C878CBD2BC CRC64;

Query Match 68.6%; Score 470; DB 11; Length 224;  
Best Local Similarity 69.9%; Pred. No. 2e-46;  
Matches 79; Conservative 16; Mismatches 18; Indels 0; Gaps 0;

QY 8 KALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYS 67  
 :||| ||| : : |||:| |||||:|||||||:|||| | |||:| |||:|  
 Db 112 RALDAAYCFRNVQDNCCLRPLYIDFKRDLGWKWIHEPKGYNANFCAGACPYLWSSDTQHT 171

QY 68 KVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 |||:| || |||:|||| | ||| |:|:| |:|||||||:| |||  
 Db 172 KVLSTYNTINPEASASPCCVSQDLEPLTILYYIGNTPKIEQLSNMIVKSRKCS 224

RESULT 13

O93449

ID O93449 PRELIMINARY; PRT; 382 AA.  
 AC O93449; Q91217;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta precursor.  
 GN TGF-BETA OR TGF.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEUKOCYTE;  
 RX MEDLINE=99242020; PubMed=10227481;  
 RA Daniels G.D., Secombes C.J.;  
 RT "Genomic organisation of rainbow trout, Oncorhynchus mykiss TGF-  
 RT BETA.";  
 RL Dev. Comp. Immunol. 23:139-147(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=LEUKOCYTE;  
 RX MEDLINE=98390168; PubMed=9722928;  
 RA Hardie L.J., Laing K.J., Daniels G.D., Grabowski P.S., Cunningham C.,  
 RA Secombes C.J.;  
 RT "Isolation of the first piscine transforming growth factor beta gene:  
 RT analysis reveals tissue specific expression and a potential regulatory  
 RT sequence in rainbow trout (Oncorhynchus mykiss).";  
 RL Cytokine 10:555-563(1998).  
 CC -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE  
 CC RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN BLOOD LEUKOCYTES, KIDNEY  
 CC MACROPHAGES, BRAIN, GILL AND SPLEEN BUT NOT IN LIVER.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AJ007836; CAA07707.1; -.  
 DR EMBL; X99303; CAA67685.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.

DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFβ.  
 DR ProDom; PD000357; TGFβ; 1.  
 DR SMART; SM00204; TGFβ; 1.  
 DR PROSITE; PS00250; TGF\_β\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT PROPEP 21 270  
 FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR β.  
 FT DISULFID 278 286 BY SIMILARITY.  
 FT DISULFID 285 348 BY SIMILARITY.  
 FT DISULFID 314 379 BY SIMILARITY.  
 FT DISULFID 318 381 BY SIMILARITY.  
 FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 237 237 N -> D (IN REF. 2).  
 FT CONFLICT 345 345 Q -> H (IN REF. 2).  
 FT CONFLICT 371 372 LS -> VP (IN REF. 2).  
 FT CONFLICT 377 377 K -> M (IN REF. 2).  
 SQ SEQUENCE 382 AA; 44136 MW; 93BD4D3540084B92 CRC64;

Query Match 68.6%; Score 470; DB 13; Length 382;  
 Best Local Similarity 75.0%; Pred. No. 3.6e-46;  
 Matches 78; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

Qy 17 SSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVIALYNQH 76  
 | ::|||:||||||| || |||:|:| |||: : :||:|||| |  
 Db 279 SDKSESCVRKLYIDFRKDLGWKWIHEPTGYFANYCIGPCTYIWNTENKYSQVLALYKHH 338  
 Qy 77 NPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||| ||||| |||||:||||: |||||:|:| |||  
 Db 339 NPGASAPCCVPQVLEPLPIIYYVGRQHKVEQLSNMIVKSCRC 382

#### RESULT 14

##### Q9PWA9

ID Q9PWA9 PRELIMINARY; PRT; 382 AA.  
 AC Q9PWA9;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta precursor.  
 GN TGF-β.  
 OS Morone chrysops x Morone saxatilis (white bass x striped bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;  
 OC Moronidae; Morone.  
 OX NCBI\_TaxID=45352;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=KIDNEY;  
 RX MEDLINE=20394636; PubMed=10938723;  
 RA Harms C.A., Kennedy-Stoskopf S., Horne W.A., Fuller F.J.,  
 RA Tompkins W.A.F.;



RT "Cloning and sequencing hybrid striped bass (*Morone saxatilis* x *M.*  
RT *chrysops*) transforming growth factor-beta (TGF-beta), and development  
RT of a reverse transcription quantitative competitive polymerase chain  
RT reaction (RT-qcPCR) assay to measure TGF-beta mRNA of teleost fish.";   
RL Fish Shellfish Immunol. 10:61-85(2000).  
CC -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE  
CC RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.  
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
CC -!- TISSUE SPECIFICITY: HIGHER LEVELS FOUND IN MONONUCLEAR CELLS FROM  
CC PERIPHERAL BLOOD THAN IN SPLEEN OR ANTERIOR KIDNEY.  
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
DR EMBL; AF140363; AAD46997.1; -.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR001839; TGFb.  
DR InterPro; IPR001111; TGFb\_N.  
DR InterPro; IPR003911; TGF\_TGFb.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFb\_propeptide; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRINTS; PR01423; TGFBETA.  
DR ProDom; PD000357; TGFb; 1.  
DR SMART; SM00204; TGFB; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein; Signal.  
FT SIGNAL 1 ? POTENTIAL.  
FT PROPEP ? 270  
FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA.  
FT DISULFID 278 286 BY SIMILARITY.  
FT DISULFID 285 348 BY SIMILARITY.  
FT DISULFID 314 379 BY SIMILARITY.  
FT DISULFID 318 381 BY SIMILARITY.  
FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT SITE 235 237 CELL ATTACHMENT SITE (POTENTIAL).  
SQ SEQUENCE 382 AA; 43846 MW; 3124D8C34EA74D72 CRC64;

Query Match 68.0%; Score 465.5; DB 13; Length 382;  
Best Local Similarity 74.5%; Pred. No. 1.2e-45;  
Matches 79; Conservative 13; Mismatches 13; Indels 1; Gaps 1;

Qy 15 CFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYN 74  
| : || ||||| ||||| : ||| : : | ||| : : ||| : |||  
Db 278 CTAQTE-TCCVRSLYIDFRKDLGWKWIHKPTGYNANYCMGSCTYIWNAENKYSQILALYK 336  
Qy 75 QHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
||||| ||||| : ||| : ||||| : |||||  
Db 337 HHNPGASAPCCVPQALEPLPIVYYVGRQHKVEQLSNMIVKSKCS 382

RESULT 15  
Q8JHF5  
ID Q8JHF5 PRELIMINARY; PRT; 379 AA.

AC Q8JHF5;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta 1.  
 OS Sparus aurata (Gilthead sea bream).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;  
 OC Sparidae; Sparus.  
 OX NCBI\_TaxID=8175;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tafalla C., Aranguren R., Secombes C.J., Castrillo J.L., Novoa B.,  
 RA Figueras A.;  
 RT "Molecular characterization of sea bream (Sparus aurata) transforming  
 RT growth factor beta1.";  
 RL Fish and Shellfish Immunol. 0:0-0(2002).  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AF510084; AAN03842.1; -.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 2.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFbeta.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 SQ SEQUENCE 379 AA; 43506 MW; C0C9D3D2FCA29C0E CRC64;

Query Match 67.8%; Score 464.5; DB 13; Length 379;  
 Best Local Similarity 74.5%; Pred. No. 1.6e-45;  
 Matches 79; Conservative 12; Mismatches 14; Indels 1; Gaps 1;

Qy 15 CFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYN 74  
 | : || |||| ||||| ||||| ||||| :| ||||:|:| | ||: : :||:| ||||  
 Db 275 CTAQTE-TCCVRSlyIDFRKDLGWKWIHKPTRYHANYCMGSCtYIWNaeNKYSQILALYK 333  
  
 Qy 75 QHNPGASAApCCVPQalePLPIVYYVGRKPKVEQLSNMIVRSCKCS 120  
 ||||| ||||| ||||| ||||| :|||: ||||| ||||| :||| ||||  
 Db 334 HHNPGASApCCVPQalePLPIlyVGRQHKVEQLSNMIVKsCKCS 379

Search completed: October 28, 2003, 09:12:24  
 Job time : 10.6208 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 10:47:27 ; Search time 4668.52 Seconds  
(without alignments)  
10489.161 Million cell updates/sec

Title: US-10-017-372E-12  
Perfect score: 1197  
Sequence: 1 atggcgcccttcggggctgcg.....gttcctgcaagtgcagctga 1197

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

```

28:  em_un:*
29:  em_vi:*
30:  em_htg_hum:*
31:  em_htg_inv:*
32:  em_htg_other:*
33:  em_htg_mus:*
34:  em_htg_pln:*
35:  em_htg_rod:*
36:  em_htg_mam:*
37:  em_htg_vrt:*
38:  em_sy:*
39:  em_htgo_hum:*
40:  em_htgo_mus:*
41:  em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	1135.8	94.9	1326	6	AX338213	AX338213 Sequence
2	1131	94.5	2221	4	AF461808	AF461808 Sus scrofa
3	1131	94.5	3206	4	PIGTGFB1A	M23703 Sus scrofa
4	1129.4	94.4	1605	4	SSTGFB1	Y00111 Porcine mRNA
5	1111.6	92.9	1750	4	GGTGFB1	X12373 Porcine mRNA
6	999.8	83.5	1173	4	OATGFB1	X76916 O.aries mRNA
7	966.2	80.7	1369	4	DOGTGFB1A	L34956 Canine tran
8	956.6	79.9	1561	9	AGMTGFB	M16658 Simian tran
9	956.6	79.9	1746	9	BC022242	BC022242 Homo sapi
10	956.6	79.9	1780	9	BC000125	BC000125 Homo sapi
11	956.6	79.9	1780	9	BC001180	BC001180 Homo sapi
12	956.6	79.9	1821	6	E03028	E03028 DNA encodin
13	956.6	79.9	2537	6	A06669	A06669 Synthetic m
14	955	79.8	2527	6	E00973	E00973 cDNA encodi
15	954.6	79.7	1173	9	BT007245	BT007245 Homo sapi
16	954.6	79.7	1173	12	BT007866	BT007866 Synthetic
17	952.4	79.6	1560	6	I06216	I06216 Sequence 2
18	952.4	79.6	1560	6	I08268	I08268 Sequence 2
19	945.4	79.0	1173	4	ECRGFB1	X99438 E.caballus
20	945.2	79.0	1176	6	AX615127	AX615127 Sequence
21	943.6	78.8	1176	6	AX481432	AX481432 Sequence
22	943.6	78.8	1176	6	AX615128	AX615128 Sequence
23	943.6	78.8	2745	9	HSTGFB1	X02812 Human mRNA
24	939	78.4	1187	4	AF175709	AF175709 Equus cab
25	936.2	78.2	1569	6	I06221	I06221 Sequence 3
26	929.4	77.6	1561	6	I08275	I08275 Sequence 3
27	889.4	74.3	1597	10	AF191297	AF191297 Cavia por
28	854.2	71.4	1585	10	RNTGFB1	X52498 Rat mRNA fo
29	847.8	70.8	1579	10	MUSTGFRNA	M13177 Mouse trans
30	847.8	70.8	1641	10	BC013738	BC013738 Mus muscu
31	847.8	70.8	2094	10	MMU009862	AJ009862 Mus muscu
32	805.6	67.3	1117	4	BOVTGFB	M36271 Bovine tran
33	804.6	67.2	1125	10	AF480858	AF480858 Sigmodon

34	682	57.0	1675	6	I03310	I03310 Sequence 1
35	654	54.6	1376	6	AX528533	AX528533 Sequence
36	652.4	54.5	1389	6	AX528619	AX528619 Sequence
37	591.4	49.4	1352	6	AX528535	AX528535 Sequence
38	586.6	49.0	1350	6	AX528615	AX528615 Sequence
39	570.6	47.7	699	6	I05434	I05434 Sequence 4
40	342.4	28.6	489	6	AX455100	AX455100 Sequence
41	342	28.6	1256	5	CHKTGFB4	M31160 Gallus gall
42	340.4	28.4	469	10	MATGFB1	X60296 M.auratus m
43	308	25.7	374	4	AF349538	AF349538 Canis fam
44	302.8	25.3	148841	9	AC011462	AC011462 Homo sapi
45	301.2	25.2	2381	9	HSTGFBG1	X05839 Human trans

# ALIGNMENTS

## RESULT 1

AX338213

LOCUS AX338213 1326 bp DNA linear PAT 09-JAN-2002

DEFINITION Sequence 1 from Patent WO0181404.

ACCESSION AX338213

VERSION AX338213.1 GI:18128750

KEYWORDS .

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

## REFERENCE

1

AUTHORS Strober,W., Nakamura,K., Kitani,A. and Fuss,I.J.

TITLE Inducible plasmid vector encoding tgf-\_g(b) and uses thereof

JOURNAL Patent: WO 0181404-A 1 01-NOV-2001;

THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)

## FEATURES

Location/Qualifiers

source

1. .1326

/organism="Sus scrofa"

/mol\_type="genomic DNA"

/db\_xref="taxon:9823"

CDS

16. .1188

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD20538.1"

/db\_xref="GI:18128751"

/translation="MAPSGLRLLPLLPLLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK  
 RIEAIRGQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD  
 YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL  
 KLKVEQHVELYQKYSNDSWRYSNRLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR  
 LSAHSSSDSKDNTLHVEINGFNSGRGDLATIHGMNRPFLLLMATPLERAQHLHSSRH  
 RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT  
 QYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS"

BASE COUNT 263 a 438 c 392 g 233 t

ORIGIN

Query Match 94.9%; Score 1135.8; DB 6; Length 1326;

Best Local Similarity 97.8%; Pred. No. 1.7e-198;

Matches 1171; Conservative 0; Mismatches 2; Indels 24; Gaps 1;

Qy	1	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	60
Db	16	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	75
Qy	61	CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	120
Db	76	CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	135
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	136	GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTCGCC	195
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	196	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	255
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCAGCCAGAG	300
Db	256	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCAGCCAGAG	315
Qy	301	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACCAAATC	360
Db	316	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACCAAATC	375
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	420
Db	376	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	435
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	436	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	495
Qy	481	AAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	540
Db	496	AAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	555
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTGATGTC	600
Db	556	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTGATGTC	615
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCTCAGT	660
Db	616	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCTCAGT	675
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	676	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	735
Qy	721	TCTGGCCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	780
Db	736	TCTGGCCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	795
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	796	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	855
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGA CTACAAGGATGACGACGACAAGGAGAAGAAC	900

```

      |||
Db      856 GATACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC 891
      |||

Qy      901 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT 960
      |||

Db      892 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT 951
      |||

Qy      961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020
      |||

Db      952 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1011
      |||

Qy      1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCGGCG 1080
      |||

Db      1012 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCGGCG 1071
      |||

Qy      1081 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
      |||

Db      1072 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1131
      |||

Qy      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAAGTGCAGCTGA 1197
      |||

Db      1132 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAAGTGCAGCTGA 1188
      |||

```

## RESULT 2

AF461808

LOCUS AF461808 2221 bp mRNA linear MAM 03-JAN-2002

DEFINITION Sus scrofa transforming growth factor beta 1 (TGFB1) mRNA, complete cds.

ACCESSION AF461808

VERSION AF461808.1 GI:18042250

KEYWORDS .

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 2221)

AUTHORS Wimmers,K., Chomdej,S., Ponsuksili,S. and Schellander,K.

TITLE Polymorphism in the porcine transforming growth factor beta 1 gene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2221)

AUTHORS Wimmers,K., Chomdej,S., Ponsuksili,S. and Schellander,K.

TITLE Direct Submission

JOURNAL Submitted (20-DEC-2001) Institute of Animal Breeding Science, University of Bonn, Endenicher Allee 15, Bonn 53115, Germany

FEATURES Location/Qualifiers

```

    source      1. .2221
                  /organism="Sus scrofa"
                  /mol_type="mRNA"
                  /db_xref="taxon:9823"
                  /chromosome="6"
    gene        1. .2221
                  /gene="TGFB1"
    CDS         1. .1173
                  /gene="TGFB1"
                  /note="cytokine"
                  /codon_start=1

```





```

Db      541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC 600
Qy      601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT 660
      |||
Db      601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT 660
Qy      661 GCCCACTCTTCCTCTGACAGCAAAGATAACA CACTCCACGTGGAAATTAACGGGTTCAAT 720
      |||
Db      661 GCCCACTGTTCCTGTGACAGCAAAGATAACA CACTCCACGTGGAAATTAACGGGTTCAAT 720
Qy      721 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
      |||
Db      721 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
Qy      781 ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 840
      |||
Db      781 ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 840
Qy      841 GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC 900
      |||
Db      841 GATACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC 876
Qy      901 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT 960
      |||
Db      877 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT 936
Qy      961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020
      |||
Db      937 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 996
Qy      1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCG 1080
      |||
Db      997 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCG 1056
Qy      1081 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
      |||
Db      1057 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1116
Qy      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTTCCTGCAAGTGCAGCTGA 1197
      |||
Db      1117 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTTCCTGCAAGTGCAGCTGA 1173

```

### RESULT 3

#### PIGTGFB1A

```

LOCUS      PIGTGFB1A                      3206 bp      mRNA      linear      MAM 31-MAR-1995
DEFINITION Sus scrofa transforming growth factor beta-1 mRNA, complete cds.
ACCESSION  M23703
VERSION    M23703.1  GI:755044
KEYWORDS   transforming growth factor-beta-1.
SOURCE     Sus scrofa (pig)
  ORGANISM Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1  (bases 1 to 3206)
  AUTHORS  Kondaiah,P., Van Obberghen-Schilling,E., Ludwig,R.L., Dhar,R.,
            Sporn,M.B. and Roberts,A.B.

```

TITLE cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.  
Evidence for alternate splicing and polyadenylation  
JOURNAL J. Biol. Chem. 263 (34), 18313-18317 (1988)  
MEDLINE 89054010  
PUBMED 2461367  
COMMENT On Apr 1, 1995 this sequence version replaced gi:341017.  
Original source text: Sus scrofa (strain miniature swine) cDNA to mRNA.

FEATURES Location/Qualifiers  
source 1. .3206  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/strain="miniature swine"  
/db\_xref="taxon:9823"  
/cell\_type="peripheral blood lymphocyte"  
gene 1. .3206  
/gene="TGF-beta-1"  
CDS 906. .2078  
/gene="TGF-beta-1"  
/codon\_start=1  
/product="transforming growth factor-beta-1"  
/protein\_id="AAA64616.1"  
/db\_xref="GI:755045"  
/translation="MPPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK  
RIEAI RQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD  
YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL  
KLKVEQHVELYQKYSNDSWRYLSNRLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR  
LSAHCSCDSKDNTLHVEINGFNSGRRGDLATIHGMNRPFLLMATPLERAQHLHSSRH  
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT  
QYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS"  
polyA\_signal 3186. .3191  
/gene="TGF-beta-1"  
polyA\_site 3206  
/gene="TGF-beta-1"  
BASE COUNT 645 a 1041 c 924 g 596 t  
ORIGIN

Query Match 94.5%; Score 1131; DB 4; Length 3206;  
Best Local Similarity 97.6%; Pred. No. 1.2e-197;  
Matches 1168; Conservative 0; Mismatches 5; Indels 24; Gaps 1;

```

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG  60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      906 ATGCCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG  965

Qy      61 CTGACGCCTTGCGCCGCCGCCGCCGACTGTCCACCTGCAAGACCATCGACATGGAGCTG  120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      966 CTGACGCCTTGCGCCGCCGCCGCCGACTGTCCACCTGCAAGACCATCGACATGGAGCTG  1025

Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC  180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     1026 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC  1085

Qy      181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT  240
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     1086 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCCGTACTGGCTCTT  1145

```

Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAAGTGTCTGAACCCGGAGCCCCGAGCCAGAG	300
Db	1146	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAAGTGTCTGAACCCGGAGCCCCGAGCCAGAG	1205
Qy	301	GCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	1206	GCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	1265
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCTGGAGCTC	420
Db	1266	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCTGGAGCTC	1325
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	1326	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	1385
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTTGGCGC	540
Db	1386	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTTGGCGC	1445
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	600
Db	1446	TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	1505
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCCTCAGT	660
Db	1506	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCCTCAGT	1565
Qy	661	GCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1566	GCCCACTGTTCTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	1625
Qy	721	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACGGCCCTTCCTGCTCCTC	780
Db	1626	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACGGCCCTTCCTGCTCCTC	1685
Qy	781	ATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGTCCCGGCACCGCCGAGCCCTG	840
Db	1686	ATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGTCCCGGCACCGCCGAGCCCTG	1745
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1746	GATACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC	1781
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	960
Db	1782	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	1841
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1842	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1901
Qy	1021	GACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCG	1080
Db	1902	GACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCG	1961
Qy	1081	GCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGCCGC	1144

```

          |||||||
Db      1962 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGCTACTACGTGGGCCGC 2021
Qy      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
          |||||||
Db      2022 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 2078

```

#### RESULT 4

##### SSTGFBR

```

LOCUS      SSTGFBR                      1605 bp      mRNA      linear      MAM 27-MAR-1995
DEFINITION Porcine mRNA for transforming growth factor-beta (TFG) precursor.
ACCESSION  Y00111
VERSION    Y00111.1  GI:2129
KEYWORDS   transforming growth factor-beta.
SOURCE     Sus scrofa (pig)
  ORGANISM Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1  (bases 1 to 1605)
  AUTHORS  Derynck,R. and Rhee,L.
  TITLE    Sequence of the porcine transforming growth factor-beta precursor
  JOURNAL  Nucleic Acids Res. 15 (7), 3187 (1987)
  MEDLINE  87174844
  PUBMED   3470708

```

```

FEATURES             Location/Qualifiers
     source            1..1605
                       /organism="Sus scrofa"
                       /mol_type="mRNA"
                       /db_xref="taxon:9823"
                       /tissue_type="ovary"
                       /clone_lib="lambda gt10"
     CDS               404..1576
                       /note="TFG-beta precursor (AA 1-390)"
                       /codon_start=1
                       /protein_id="CAA68291.1"
                       /db_xref="GI:2130"
                       /db_xref="SWISS-PROT:P07200"
                       /translation="MPPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK
R IEAIRGQILSKRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD
YYAKEVTRVLMLESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNDSWRYLNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR
LSAHCSCDSKDNTLHVEINGFNSGRGDLATIHGMNRPFLLMATPLERAQHLHSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"
     mat_peptide       1238..1573
                       /product="mature TFG-beta"
     misc_feature      647..655
                       /note="pot. N-glycosylation site"
     misc_feature      809..817
                       /note="pot. N-glycosylation site"
BASE COUNT      298 a      572 c      437 g      298 t
ORIGIN

```

```

Query Match          94.4%;  Score 1129.4;  DB 4;  Length 1605;
Best Local Similarity 97.5%;  Pred. No. 2.6e-197;
Matches 1167;  Conservative 0;  Mismatches 6;  Indels 24;  Gaps 1;

```

Qy	1	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	60
Db	404	ATGCCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	463
Qy	61	CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	120
Db	464	CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	523
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	524	GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGCCAGATTCTGTCCAAGCTTCGGCTCGCC	583
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	584	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCCGTACTGGCTCTT	643
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCCGAGCCAGAG	300
Db	644	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCCGAGCCAGAG	703
Qy	301	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	704	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGCTGGAAAGCGGCAACCAAATC	763
Qy	361	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGGAGCTC	420
Db	764	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGGAGCTC	823
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	824	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	883
Qy	481	AAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	540
Db	884	AAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	943
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCCGAGTGGCTGTCTTTGATGTC	600
Db	944	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCCGAGTGGCTGTCTTTGATGTC	1003
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCTCAGT	660
Db	1004	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCTCAGT	1063
Qy	661	GCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1064	GCCCACTGTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	1123
Qy	721	TCTGGCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	780
Db	1124	TCTGGCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	1183
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1184	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	1243

```

Qy      841 GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC 900
      |||
Db      1244 GATACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC 1279

Qy      901 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT 960
      |||
Db      1280 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT 1339

Qy      961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020
      |||
Db      1340 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1399

Qy     1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCG 1080
      |||
Db     1400 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCG 1459

Qy     1081 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
      |||
Db     1460 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1519

Qy     1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAAGTGCAGCTGA 1197
      |||
Db     1520 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAAGTGCAGCTGA 1576

```

# RESULT 5

GGTGFB1

LOCUS GGTGFB1 1750 bp mRNA linear MAM 27-MAR-1996

DEFINITION Porcine mRNA for transforming growth factor-beta 1.

ACCESSION X12373

VERSION X12373.1 GI:63808

KEYWORDS transforming growth factor-beta 1.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 1750)

AUTHORS Jakowlew,S.B., Dillard,P.J., Sporn,M.B. and Roberts,A.B.

TITLE Nucleotide sequence of chicken transforming growth factor-beta 1  
(TGF-beta 1)

JOURNAL Nucleic Acids Res. 16 (17), 8730 (1988)

MEDLINE 88335639

PUBMED 3166520

REFERENCE 2 (bases 1 to 1750)

AUTHORS Jakowlew,S.B.

TITLE Direct Submission

JOURNAL Submitted (14-JUL-1988) Jakowlew S.B., National Institute of  
health, National Cancer Institute, Laboratory of Chemoprevention,  
Building 41, Room B902, Bethesda, Maryland 20892, USA

COMMENT The submitters believe that the chicken cDNA library was  
contaminated with porcine cDNA, and that the sequence is infact  
porcine TGF-beta-1. 27-MAR-1996.

FEATURES Location/Qualifiers

source 1..1750

/organism="Sus scrofa"

/mol\_type="mRNA"

/strain="white leghorn"

```

        /db_xref="taxon:9823"
        /clone="pTGFB-ChX119"
        /cell_type="chondrocyte"
gene      1. .1750
        /gene="TGF-beta 1"
5'UTR     1. .446
        /gene="TGF-beta 1"
CDS       447. .1622
        /gene="TGF-beta 1"
        /codon_start=1
        /product="transforming growth factor"
        /protein_id="CAA30933.1"
        /db_xref="GI:63809"
        /db_xref="SWISS-PROT:P07200"
        /translation="MPPSGPGLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK
RIEAIRGQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD
YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNDSWGYLSNRLAPSDSPEWLSFDVTGVVRQWLTRREAI EGFR
LSAHCSCDSKDNTLHVEINAGFNSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSR
HRRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD
TQYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS"
mat_peptide 1285. .1619
        /gene="TGF-beta 1"
3'UTR     1620. .1750
        /gene="TGF-beta 1"
BASE COUNT      325 a      627 c      479 g      319 t
ORIGIN

```

```

Query Match          92.9%;  Score 1111.6;  DB 4;  Length 1750;
Best Local Similarity 97.0%;  Pred. No. 4.7e-194;
Matches 1164;  Conservative 0;  Mismatches 9;  Indels 27;  Gaps 2;

```

```

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| ||||||||||| |||||||||||||||||||||||||||||||||||
Db      447 ATGCCGCCTTCGGGGCCTGGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 506

Qy      61 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||||||||||| |||||||||||||||||||||||||||||||||||
Db      507 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 566

Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||||||||||| |||||||||||||||||||||||||||||||||||
Db      567 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTCGCC 626

Qy      181 AGCCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||||||||||| |||||||||||||||||||||||||||||||||||
Db      627 AGCCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCCGTACTGGCTCTT 686

Qy      241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCCGAGCCAGAG 300
      ||||||||||| |||||||||||||||||||||||||||||||||||
Db      687 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCCGAGCCAGAG 746

Qy      301 GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      ||||||||||| |||||||||||||||||||||||||||||||||||
Db      747 GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 806

Qy      361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420

```

Db	807	 TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	866
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	867	 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	926
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAATACAGCAATGATTCTGGCGC	540
Db	927	 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAATACAGCAATGATTCTGGGGC	986
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC	600
Db	987	 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC	1046
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	1047	 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	1106
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAAC---GGGTTC	717
Db	1107	 GCCCACTGTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGCAGGGTTC	1166
Qy	718	AATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTC	777
Db	1167	 AATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTC	1226
Qy	778	CTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC	837
Db	1227	 CTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC	1286
Qy	838	CTGGATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAG	897
Db	1287	 CTGGATACCAACTACTGCTTCAGCTCCAC-----GGAGAAG	1322
Qy	898	AACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATT	957
Db	1323	 AACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATT	1382
Qy	958	CATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGC	1017
Db	1383	 CATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGC	1442
Qy	1018	CTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCG	1077
Db	1443	 CTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCG	1502
Qy	1078	GCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGC	1137
Db	1503	 GCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGC	1562
Qy	1138	CGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAAGTGCAGCTGA	1197
Db	1563	 CGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAAGTGCAGCTGA	1622





Db	61	CTGACGCCTGGCCGGCCGGTCGCCGGA	CTGTCCACCTGCAAGACCATCGACATGGAGCTG	120
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATT	CGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	121	GTGAAGCGGAAGGGCATCGAGGCCATCCGCGGT	CAGATTTTGTCCAACTTCGGCTCGCC	180
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT		240
Db	181	AGTCCCCCGAGCCAGGGGGACGTGCCACCCGGCCCGCTGCCCGAGGCCATACTGGCCCTT		240
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCAGCCAGAG		300
Db	241	TACAACAGTACCCGCGACCGGGTGGCCGGGGAAAGTGCCGAAACGGAGCCTGAGCCAGAG		300
Qy	301	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC		360
Db	301	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAATACGGCAACCAAATC		360
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC		420
Db	361	TATGACAAAATGAAGTCTAGCTCGCACAGCATATATATGTTCTTCAACACGTCGGAGCTC		420
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC		480
Db	421	CGGGAAGCAGTGCTGAACCTGTGTTGCTCTCTCGGGCAGACGTGCGCCTGCTGAGGCTC		480
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTTGGCGC		540
Db	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATATAGCAACAATTCTTGGCGC		540
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA	CTACCCGGAGTGGCTGTCTTTGATGTC	600
Db	541	TACCTCAGCAACCGGCTGCTGCCCCCAGCGACTACCCGGAGTGGCTGTCTTTGACGTC		600
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTCGCCTCAGT		660
Db	601	ACTGGAGTTGTGCGGCAGTGGCTGACCCACAGAGAGGAAATAGAAGGCTTTCGCCTCAGT		660
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT		720
Db	661	GCCCACTGTTCCTGTGACAGTAAGGATAACACGCTTCAAGTGGACATCAACGGGTTCAAT		720
Qy	721	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC		780
Db	721	TCCGGCCGCCGGGGTGACCTGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC		780
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG		840
Db	781	ATGGCCACCCCTCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG		840
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC		900
Db	841	GACACCAACTACTGCTTCAGCTCCAC-----AGAAAAGAAC		876
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT		960
Db	877	TGCTGTGTTCTGTCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAC		936

```

Qy      961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020
      |||
Db      937 GAACCCAAGGGCTACCACGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTG 996

Qy      1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCGGCG 1080
      |||
Db      997 GACACACAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCACAAACCCGGGCGCATCGGCG 1056

Qy      1081 GCGCCGTGCTGCGTGCCGCAAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
      |||
Db      1057 GCGCCGTGCTGCGTGCCCTCAGGCGCTGGAACCCCTGCCCATCGTGTACTACGTGGGCCGC 1116

Qy      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
      |||
Db      1117 AAGCCCAAGGTGGAGCAGTTGTCCAACATGATCGTGCGCTCCTGCAAGTGCAGCTGA 1173

```

# RESULT 7

DOGTGFB1A

LOCUS DOGTGFB1A 1369 bp mRNA linear MAM 30-OCT-1994

DEFINITION Canine transforming growth factor-beta 1 (TGFB1) mRNA, complete cds.

ACCESSION L34956

VERSION L34956.1 GI:516071

KEYWORDS homologue; transforming growth factor-beta 1.

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 1369)

AUTHORS Manning,A.M., Auchampach,J.A., Drong,R.F. and Slightom,J.L.

TITLE Cloning of a canine cDNA homologous to human transforming growth factor-beta 1 (TGFBeta1)

JOURNAL Unpublished (1994)

COMMENT Original source text: Canis familiaris adult jugular vein endothelial cDNA to mRNA.

FEATURES Location/Qualifiers

source 1..1369  
 /organism="Canis familiaris"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9615"  
 /cell\_type="LPS-activated"  
 /tissue\_type="jugular vein endothelial"  
 /dev\_stage="adult"

gene 1..1369  
 /gene="TGFB1"

5'UTR 1..57  
 /gene="TGFB1"

CDS 58..1230  
 /gene="TGFB1"  
 /function="anti-inflammatory agent"  
 /note="precursor"  
 /codon\_start=1  
 /product="transforming growth factor-beta 1"  
 /protein\_id="AAA51458.1"  
 /db\_xref="GI:516072"

```

/translation="MPPSGLRLLPLLLPLLRLLVLTPGRPAAGLSTCKTIDMELVKRK
RIEAIRGQILSKLRLSSPPSQGEVPPVPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
YYAKEVTRVLMVENTNKIYEKVKKSPHSIYMLFNTSELREAVPEPVLLSRAELRLLRL
KLKAEQHVELYQKYSNDSWRYLNRLLAPSDTPEWLSFDVTGVVRQWLSHGGEVEGFR
LSAHCSCDSKDNLTQVDINGFSSRRGDLATI HGMNRPFLLL MATPLERAQHLHSSRQ
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS"
mat_peptide 1039. .1227
/gene="TGFB1"
/product="transforming growth factor-beta 1"
3'UTR 1231. .1369
/gene="TGFB1"
BASE COUNT 264 a 473 c 415 g 216 t 1 others
ORIGIN
Query Match 80.7%; Score 966.2; DB 4; Length 1369;
Best Local Similarity 89.0%; Pred. No. 2.3e-167;
Matches 1065; Conservative 0; Mismatches 108; Indels 24; Gaps 1;

Qy 1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
||| |||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 58 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTGCCGCTGCTGCGGCTACTAGTG 117

Qy 61 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 118 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAACTG 177

Qy 121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 178 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCTCC 237

Qy 181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 238 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGTGCCGCTGCCCGAGGCCGTGCTGGCCCTC 297

Qy 241 TACAACAGTACCCGCGACCGGGTAGCCGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 298 TACAACAGCACCCGCGACCGGGTGGCGGGGAGAGCGCCGAGCCGAGCCCGAGCCCGAG 357

Qy 301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 358 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAACACCAACAAAATC 417

Qy 361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 418 TATGAGAAAAGTCAAGAAAAGTCCGCACAGCATATATATGCTCTTCAACACATCAGAGCTC 477

Qy 421 CGGGAAGCGGTGCCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 478 CGAGAAGCAGTGCTGAGCCCGTCTTGCTCTCCCGGCAGAGTTGCGCCTGCTGAGGCTC 537

Qy 481 AAGTTAAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC 540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 538 AAGTTAAAAAGCGGAGCAGCATGTGGAGCTGTACCAGAAATATAGCAATGATTCCTGGCGC 597

Qy 541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCCGAGTGGCTGTCCTTTGATGTC 600
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db      598 TACCTCAGCAACCGGCTGCTGGCGCCCAGCGACACGCCAGAATGGCTGTCTTTGATGTC 657
Qy      601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT 660
      || ||||| ||| ||||| ||||| ||| ||| || ||||| ||||| |||||
Db      658 ACTGGAGTCGTGAGGCAGTGGCTGAGCCATGGAGGGGAAGTCGAGGGCTTTTCGCCTCAGT 717
Qy      661 GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720
      ||||| ||||| ||||| ||||| ||||| ||| ||| || ||||| ||||| ||
Db      718 GCCCACTGTTCTCTGTGACAGCAAAGATAACACACTGCAAGTAGACATTAACGGGTTCAAT 777
Qy      721 TCTGGCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db      778 TCCAGCCGCGAGGTGACCTGGCCACCATTATGGCATGAACCGACCTTCCTGCTCCTC 837
Qy      781 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 840
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      838 ATGGCCACCCCACTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCAGCGCCGGGCCCTG 897
Qy      841 GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC 900
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      898 GACACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC 933
Qy      901 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT 960
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      934 TGCTGCGTCCGGCAGCTCTACATTGACTTCCGCAAGGATCTGGGCTGGAAGTGGATCCAT 993
Qy      961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      994 GAGCCCAAGGGTTACCACGCTAACTTCTGCCTGGGGCCCTGCCCCTACATTTGGAGCCTG 1053
Qy      1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCG 1080
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1054 GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCACAAACCGGGCGCGTCGGCG 1113
Qy      1081 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1114 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1173
Qy      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1174 AAGCCCAAGGTGGAGCAGCTGTGAACATGATCGTGCGCTCCTGCAAGTGCAGCTGA 1230

```

# RESULT 8

AGMTGFB

LOCUS AGMTGFB 1561 bp mRNA linear PRI 27-APR-1993

DEFINITION Simian transforming growth factor-beta (TGF) mRNA, complete cds.

ACCESSION M16658

VERSION M16658.1 GI:176552

KEYWORDS growth factor; transforming growth factor-beta.

SOURCE Cercopithecus aethiops (African green monkey)

ORGANISM Cercopithecus aethiops

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
Cercopithecinae; Cercopithecus.

REFERENCE 1 (bases 1 to 1561)

AUTHORS Sharples,K., Plowman,G.D., Rose,T.M., Twardzik,D.R. and

Purchio,A.F.  
 TITLE Cloning and sequence analysis of simian transforming growth factor-beta cDNA  
 JOURNAL DNA 6 (3), 239-244 (1987)  
 MEDLINE 87246074  
 PUBMED 3474130  
 COMMENT Original source text: African green monkey cells (cell line BSC-40), cDNA to mRNA, clone pTGF-beta-2.  
 FEATURES Location/Qualifiers  
   source 1..1561  
           /organism="Cercopithecus aethiops"  
           /mol\_type="mRNA"  
           /db\_xref="taxon:9534"  
   CDS 262..1434  
       /note="transforming growth factor-beta precursor"  
       /codon\_start=1  
       /protein\_id="AAA35369.1"  
       /db\_xref="GI:176553"  
       /translation="MPPSGLRLLPLLLPLLWLLVLTSPRPAAGLSTCKTIDMELVKRKRIETIRGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWRYLSNRLLAPSNSPEWLSFDVTGVVRQWLRSRGGEIEGFRLSAHCSCDSKDNLTQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRRRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS"  
   sig\_peptide 262..1095  
               /note="transforming growth factor-beta signal peptide"  
   mat\_peptide 1096..1431  
               /product="transforming growth factor-beta"  
 BASE COUNT 301 a 547 c 446 g 267 t  
 ORIGIN

Query Match 79.9%; Score 956.6; DB 9; Length 1561;  
 Best Local Similarity 88.5%; Pred. No. 1.3e-165;  
 Matches 1059; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

```

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      262 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 321

Qy      61 CTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      322 CTGACGCCTAGCCGGCCGGCCGGCCGAGACTATCCACCTGCAAGACTATCGACATGGAGCTG 381

Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      382 GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 441

Qy      181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      442 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCGGTGCTCGCCCTG 501

Qy      241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGCCAGAG 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      502 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCGGAGCCGGAGCCCGAACCGGAG 561

Qy      301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACCAAATC 360

```

Database	Accession	Sequence	Length
Db	562	GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACGAAATC	621
Qy	361	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	420
Db	622	TATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	681
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	682	CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	741
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTTGGCGC	540
Db	742	AAGTTAAAAGTGGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTTGGCGA	801
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	600
Db	802	TACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	861
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	862	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	921
Qy	661	GCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	922	GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT	981
Qy	721	TCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTTCCTGCTCCTC	780
Db	982	ACCGGCCGCCGAGGTGACCTGGCCACAATTCAATGGCATGAACCGGCCCTTTCCTGCTTCTC	1041
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGTCTCCGGCACCGCCGAGCCCTG	840
Db	1042	ATGGCCACCCCACTGGAGAGGGCCCAACATCTGCAAAGTCTCCGGCACCGCCGAGCCCTG	1101
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1102	GACACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC	1137
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT	960
Db	1138	TGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1197
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1198	GAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATTTGGAGCCTG	1257
Qy	1021	GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCGGCG	1080
Db	1258	GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCG	1317
Qy	1081	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGGCCGC	1140
Db	1318	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGGCCGC	1377
Qy	1141	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAAGTGCAGCTGA	1197

Db 1378 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTCCTGCAAATGCAGCTGA 1434

RESULT 9

BC022242

LOCUS BC022242 1746 bp mRNA linear PRI 04-FEB-2002

DEFINITION Homo sapiens, clone MGC:22008 IMAGE:4399762, mRNA, complete cds.

ACCESSION BC022242

VERSION BC022242.1 GI:18490115

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1746)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (01-FEB-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 27 Row: e Column: 21

This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis.

FEATURES Location/Qualifiers

source

1..1746

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:22008 IMAGE:4399762"

/tissue\_type="Duodenum, adenocarcinoma"

/clone\_lib="NIH\_MGC\_88"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

CDS

370..1542

/codon\_start=1

/product="Unknown (protein for MGC:22008)"

/protein\_id="AAH22242.1"

/db\_xref="GI:18490116"

/translation="MPPSGLRLLLLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK  
RIEAI RQILSKLR LASPPSQGEVPPGP LPEAVLALYNSTRDRVAGESAEPEPEPEAD



YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRL  
 KLKVEQHVELYQKYSNNSWRYLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFR  
 LSAHCSCDSRDNTLQVDINGFTTGRRGDLATI HGMNRPFLLLMATPLERAQHLQSSRH  
 RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT  
 QYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"

BASE COUNT        376 a        612 c        472 g        286 t  
 ORIGIN

Query Match                79.9%;    Score 956.6;    DB 9;    Length 1746;  
 Best Local Similarity    88.5%;    Pred. No. 1.3e-165;  
 Matches 1059;    Conservative    0;    Mismatches 114;    Indels    24;    Gaps    1;

Qy	1	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGGCTGCTGTGGCTGCTAGTG	60
Db	370	ATGGCGCCTTCGGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTACTGGTG	429
Qy	61	CTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	120
Db	430	CTGACGCCTGGCCGGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	489
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	490	GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	549
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	550	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCCGAGGCCGTGCTCGCCCTG	609
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCCGAGCCAGAG	300
Db	610	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCCGAGCCTGAG	669
Qy	301	GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	670	GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC	729
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	420
Db	730	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	789
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	790	CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	849
Qy	481	AAGTTAAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	540
Db	850	AAGTTAAAAAGTGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	909
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTGATGTC	600
Db	910	TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTGATGTC	969
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCG CAGAGAGGCTATAGAGGGTTTTCGCCTCAGT	660
Db	970	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTCGCCTTAGC	1029
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720

Accession	Position	Sequence	Position
Db	1030	GCCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT	1089
Qy	721	TCTGGCCGCGCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCTGCTCCTC	780
Db	1090	ACCGGCCGCGCGAGGTGACCTGGCCACCATTGATGGCATGAACCGGCCCTTCTGCTTCTC	1149
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1150	ATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1209
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1210	GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC	1245
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTGAT	960
Db	1246	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1305
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1306	GAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCTACATTTGGAGCCTG	1365
Qy	1021	GACACTCAGTACAGCAAGGTCTGTGCTCTGTACAACCAGCACAACCCGGGCGCGTTCGGCG	1080
Db	1366	GACACGCAGTACAGCAAGGTCTGTGCTCTGTACAACCAGCATAACCCGGGCGCGTTCGGCG	1425
Qy	1081	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGCG	1140
Db	1426	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGCG	1485
Qy	1141	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTTCTGCAAGTGCAGCTGA	1197
Db	1486	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTTCTGCAAGTGCAGCTGA	1542

RESULT 10

BC000125

LOCUS BC000125 1780 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, Similar to transforming growth factor, beta 1, clone MGC:3119 IMAGE:3351664, mRNA, complete cds.

ACCESSION BC000125

VERSION BC000125.1 GI:12652748

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1780)

AUTHORS      Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (03-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT           Contact: MGC help desk



Db	507	CTGACGCCTGGCCGCGCCGCGCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	566
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	567	GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	626
Qy	181	AGCCCCCGGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	627	AGCCCCCGGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG	686
Qy	241	TACAAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG	300
Db	687	TACAAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG	746
Qy	301	GCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACCAAATC	360
Db	747	GCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCCACAACGAAATC	806
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTC	420
Db	807	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	866
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	867	CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	926
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC	540
Db	927	AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTTGGCGA	986
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	600
Db	987	TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTGATGTC	1046
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	1047	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	1106
Qy	661	GCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1107	GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT	1166
Qy	721	TCTGGCCGCCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTC	780
Db	1167	ACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGCTTCTC	1226
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1227	ATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1286
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1287	GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC	1322
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT	960
Db	1323	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1382

Qy 961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020  
 || |||||  
 Db 1383 GAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTACATTTGGAGCCTG 1442  
 Qy 1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCG 1080  
 |||||  
 Db 1443 GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCGGGCGCCTCGGCG 1502  
 Qy 1081 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140  
 |||||  
 Db 1503 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC 1562  
 Qy 1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197  
 |||||  
 Db 1563 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAAGTGCAGCTGA 1619

RESULT 11

BC001180

LOCUS BC001180 1780 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, Similar to transforming growth factor, beta 1, clone  
 MGC:2323 IMAGE:3356605, mRNA, complete cds.

ACCESSION BC001180

VERSION BC001180.1 GI:12654682

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1780)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (11-DEC-2000) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline  
 Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,  
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 6 Row: e Column: 10  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 37097.

```

FEATURES             Location/Qualifiers
     source            1..1780
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="MGC:2323 IMAGE:3356605"
                        /tissue_type="Eye, retinoblastoma"
                        /clone_lib="NIH_MGC_16"
                        /lab_host="DH10B-R"
                        /note="Vector: pOTB7"
     CDS               447..1619
                        /codon_start=1
                        /product="Similar to transforming growth factor, beta 1"
                        /protein_id="AAH01180.1"
                        /db_xref="GI:12654683"
                        /translation="MPPSGLRLLLLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK
R IEAIRGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNNSWRYL SNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFR
LSAHCSCDSRDNTLQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"

```

BASE COUNT        341 a        648 c        493 g        298 t  
 ORIGIN

Query Match                    79.9%;    Score 956.6;    DB 9;    Length 1780;  
 Best Local Similarity        88.5%;    Pred. No. 1.3e-165;  
 Matches 1059;    Conservative    0;    Mismatches 114;    Indels    24;    Gaps    1;

```

Qy      1  ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG  60
      ||| |||| || ||||||||| ||| ||||||||| ||||||||| ||| |||
Db     447  ATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTACTGGTG  506

Qy      61  CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG  120
      ||||||||| ||||||||| ||||| ||||||||| ||||||||| ||| |||
Db     507  CTGACGCCTGGCCGGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG  566

Qy     121  GTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC  180
      ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||| |||
Db     567  GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC  626

Qy     181  AGCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT  240
      ||||||||| ||||||||| ||||||||| ||||||||| ||| ||| |||
Db     627  AGCCCCCGAGCCAGGGGGAGGTGCCGCCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG  686

Qy     241  TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCAGCCAGAG  300
      ||||||||| ||||||||| ||||||||| ||||| ||||||||| ||||| |||
Db     687  TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCAGCCTGAG  746

Qy     301  GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACCAAATC  360
      || ||||||||| ||||||||| ||||||||| ||||||||| ||| ||| |||
Db     747  GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACGAAATC  806

Qy     361  TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC  420

```

Db	807	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	866
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	867	CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	926
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	540
Db	927	AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA	986
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC	600
Db	987	TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTGATGTC	1046
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	1047	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	1106
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1107	GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT	1166
Qy	721	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	780
Db	1167	ACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTC	1226
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1227	ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1286
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGAATAAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1287	GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC	1322
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT	960
Db	1323	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1382
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1383	GAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCCTG	1442
Qy	1021	GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCG	1080
Db	1443	GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCG	1502
Qy	1081	GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1140
Db	1503	GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC	1562
Qy	1141	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAAGTGCAGCTGA	1197
Db	1563	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCTTGCAAGTGCAGCTGA	1619

E03028  
 LOCUS E03028 1821 bp RNA linear PAT 29-SEP-1997  
 DEFINITION DNA encoding human prepro TGF-beta1.  
 ACCESSION E03028  
 VERSION E03028.1 GI:2171250  
 KEYWORDS JP 1991180192-A/1.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1821)  
 AUTHORS Ohashi,H., Ishii,Y., Miyata,Y., Miyazono,K., Miyagawa,K. and  
 Takaku,F.  
 TITLE PRODUCTION OF HUMAN PRO-TGF-BETA1 BY GENETIC RECOMBINATION  
 JOURNAL Patent: JP 1991180192-A 1 06-AUG-1991;  
 KIRIN BREWERY CO LTD  
 COMMENT OS Homo sapiens (human)  
 PN JP 1991180192-A/1  
 PD 06-AUG-1991  
 PF 07-DEC-1989 JP 1989318243  
 PI OHASHI HIDEYA, ISHII YASUYUKI, MIYATA YOSHINORI, PI MIYAZONO  
 KOHEI,  
 PI MIYAGAWA KIYOSHI, TAKAKU FUMIMARO  
 PC C12P21/00,C07K13/00,C12N5/10,C12N15/18//C12Q1/68,(C12P21/00,  
 PC C12R1:91),  
 PC (C12N15/18,C12R1:91);  
 CC strandedness: Double;  
 CC topology: Linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 CC \*source: tissue\_type=placenta;  
 CC \*source: clone=pVC19-TGF;  
 FH Key Location/Qualifiers  
 FT 5'UTR 1. .510  
 FT CDS 511. .1683  
 FT /product='human prepro TGF-beta1' FT  
 sig\_peptide 511. .597  
 FT mat\_peptide 598. .1680  
 FT /note='human pro TGF-beta1'  
 FT mat\_peptide 1345. .1680  
 FT /product='human TGF-beta1'  
 FT 3'UTR 1684. .1820.  
 FEATURES Location/Qualifiers  
 source 1. .1821  
 /organism="Homo sapiens"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 326 a 679 c 508 g 308 t  
 ORIGIN

Query Match 79.9%; Score 956.6; DB 6; Length 1821;  
 Best Local Similarity 88.5%; Pred. No. 1.3e-165;  
 Matches 1059; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

Qy 1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60  
 ||| |||| || ||||||||| ||||||||||| ||||||||||| ||||



D  
Q  
  
D  
Q  
  
D  
Q  
  
D  
Q  
  
D  
Q  
  
D  
Q  
  
D  
Q  
  
D  
Q  
  
D  
Q  
  
D  
Q  
  
D  
Q  
  
D

512 ATGCCGCGCTCCGGGCTGCGGCTGCTGCCGTGCTGCTACCGTGCTGTGGCTACTGGT 571

61 CTGACGCCTGGCCGCGC GCGCCG CCGG ACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120  
|||||  
572 CTGACGCCTGGCCGCGC GCGCCG CCGG ACTATCCACCTGCAAGACTATCGACATGGAGCTG 631

121 GTGAAGCGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180  
|||||  
632 GTGAAGCGAAGCGCATCGAGGCCATTCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 691

181 AGCCCCCGAGCCAGGGGGACGTGCCGCCGCGCCCGCTGCCTGAGGCAGTA TGGCTCTT 240  
|||||  
692 AGCCCCCGAGCCAGGGGGAGGTGCCGCCGCGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 751

241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTTGTCGAACCGGAGCCCAGCCAGAG 300  
|||||  
752 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCAGCCTGAG 811

301 GCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGAAAAGCGGCAACCAAATC 360  
|||  
812 GCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGAAAA CCACAACGAAATC 871

361 TATGATAAAATTCAAGGGCACCCCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420  
||||| ||  
872 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 931

421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480  
|| |  
932 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGAGGCTC 991

481 AAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC 540  
|||||  
992 AAGTTAAAAGTGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA 1051

541 TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGCTGTCTTTGATGTC 600  
|||||  
1052 TACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGTTATCTTTGATGTC 1111

601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGT 660  
|||||  
1112 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCTTAGC 1171

661 GCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720  
|||||  
1172 GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGACATCAACGGGTTCACT 1231

721 TCTGGCCGCGCGGGGTGACCTGGCCACCATTACAGGCATGAACCGGCCCTTCCTGCTCCTC 780  
| |  
1232 ACCGGCCGCGGAGGTGACCTGGCCACCATT CATGGCATGAACCGGCCCTTCCTGCTTCTC 1291

781 ATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGTCCCGGCACCGCCGAGCCCTG 840  
|||||  
1292 ATGGCCACCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG 1351

841 GATACCAACTACTGCTTCAGCTCCACGACTACAAGGATGACGACGACAAGGAGAAGA AC 900  
|| |  
1352 GACACCAACTATTGCTTCAGCTCCAC----- GGAGAAGA AC 1387

Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	960
Db	1388	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1447
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1448	GAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCTACATTTGGAGCCTG	1507
Qy	1021	GACACTCAGTACAGCAAGGTCTTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCTGGCG	1080
Db	1508	GACACGCAGTACAGCAAGGTCTTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCTGGCG	1567
Qy	1081	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1140
Db	1568	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC	1627
Qy	1141	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1197
Db	1628	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1684

## RESULT 13

LOCUS

DEFINITION Synthetic mRNA for preTGF-Betal.

ACCESSION A06669

VERSION A06669

KEYWORDS

KEYWORDS  
SOURCE

SOURCE	syncretic construct
ORGANISM	synthetic construct

ORGANISM    synthetic construct  
              artificial sequences

REFERENCE 1. (pages 1 to 3537)

## REFERENCE

## AUTHORS

AUTHORS .  
TITLE NUCLEIC ACID ENCODING TCE beta-3 AND ITS USE

JOURNAL Patent, HQ 9913101, A 4-14, FEB 1999

JOURNAL Patent: WO 8912101-A 4 14-DE  
FEATURES Location/Qualification

```
FEATURES      Location/Qualifiers
    source     1..2537
               /organism="synthetic construct"
```

```

/mol_type="mRNA"
/db_xref="taxon:32630"

```

```

842: 2014
/codon_start=1
/transl_table=11
/product="preTGF-beta1"
/protein_id="CAA00588.1"
/db_xref="GI:412941"
/translation="MPPSGLRLLPLLLPLLWLLVLTGPPAAGLSTCKTIDMELVKRK
RIEAIRGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNNSWRYLSNRLLAPSDSPEWLSFDVTGTVVRQWLSRGGEI EGFR
LSAHCSDNDRNTLQVDINGFTTGRRGDLATI HGMNRPVLLMATPLERAQHLQSSRH
RRALDTCNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"

```

BASE COUNT            473 a       893 c       739 g       432 t

ORIGIN

Query Match 79.9%; Score 956.6; DB 6; Length 2537;  
Best Local Similarity 88.5%; Pred. No. 1.2e-165;  
Matches 1059; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

```
Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    842 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 901

Qy     61 CTGACGCCTGGCCGCGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    902 CTGACGCCTGGCCGCGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 961

Qy    121 GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    962 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 1021

Qy    181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1022 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 1081

Qy    241 TACAACAGTACCCGCGACCGGGTAGCCGGGAAAAGTGTGCAACCGGAGCCCGAGCCAGAG 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1082 TACAACAGCACCCGCGACCGGGTGGCCGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 1141

Qy    301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1142 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC 1201

Qy    361 TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGCGAGCTC 420
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1202 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 1261

Qy    421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1262 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 1321

Qy    481 AAGTTAAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAATACAGCAATGATTCCTGGCGC 540
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1322 AAGTTAAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA 1381

Qy    541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTGATGTC 600
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1382 TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTGATGTC 1441

Qy    601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT 660
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1442 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC 1501

Qy    661 GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1502 GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT 1561

Qy    721 TCTGGCCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1562 ACCGGCCCGCAGGTTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTC 1621

Qy    781 ATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 840
```

Db	1622		ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1681
Qy	841		GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1682		GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC	1717
Qy	901		TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	960
Db	1718		TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1777
Qy	961		GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1778		GAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTACATTTGGAGCCTG	1837
Qy	1021		GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCCGGCG	1080
Db	1838		GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCCGGCG	1897
Qy	1081		GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCCGC	1140
Db	1898		GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCCGC	1957
Qy	1141		AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1197
Db	1958		AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	2014

# RESULT 14

E00973

LOCUS E00973 2527 bp RNA linear PAT 29-SEP-1997

DEFINITION cDNA encoding human TGF-beta.

ACCESSION E00973

VERSION E00973.1 GI:2169234

KEYWORDS JP 1986219395-A/1.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2527)

AUTHORS Riku,M.A.D.D. and Debitsudo,B.G.

TITLE NUCLEIC ACID CODED WITH TGF-BETA AND ITS USE

JOURNAL Patent: JP 1986219395-A 1 29-SEP-1986;

GENENTECH INC

COMMENT

OS human

PN JP 1986219395-A/1

PD 29-SEP-1986

PF 20-MAR-1986 JP 1986064661

PR 22-MAR-1985 US 85 715142

PI RIKU MAIKERU ANDORE DERINKU, DEBITSUDO BANNOOMAN GETSUDERU PC  
C12P21/00,C12N1/00,C12N5/00,C12N15/00//C12Q1/68,(C12P21/00, PC  
C12R1:91),

PC (C12N1/00,C12R1:19),(C12N5/00,C12R1:91),(C12N15/00,C12R1:91);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

```

CC      *source: tissue_type=placenta and glyoblastoma; FH      Key
        Location/Qualifiers
FH
FT      CDS                      842..2014
FT                      /product='pre TGF-beta'
FT      mat_peptide              1676..2011
FT                      /product='TGF-beta'
FT      5'UTR                   1..841
FT      3'UTR                   2015..2537
FT      stem_loop               37..113
FT      stem_loop               2015..2100
FT      polyA_site              2514..2519.
FEATURES             Location/Qualifiers
     source                1..2527
                           /organism="Homo sapiens"
                           /mol_type="genomic RNA"
                           /db_xref="taxon:9606"
BASE COUNT      472 a      888 c      735 g      432 t
ORIGIN

Query Match          79.8%;  Score 955;  DB 6;  Length 2527;
Best Local Similarity 88.4%;  Pred. No. 2.4e-165;
Matches 1058;  Conservative 0;  Mismatches 115;  Indels 24;  Gaps 1;

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      842 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 901

Qy      61 CTGACGCCTGGCCCGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      902 CTGACGCCTGGCCCGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 961

Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      962 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 1021

Qy      181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1022 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTC 1081

Qy      241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGCCAGAG 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1082 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 1141

Qy      301 GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      || || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1142 GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC 1201

Qy      361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420
      ||||| || ||||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1202 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 1261

Qy      421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
      || ||||| ||| ||||| || ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
Db      1262 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 1321

Qy      481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC 540

```

```

      |||
Db    1322 AAGTTAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA 1381
      |||
Qy    541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTGATGTC 600
      |||
Db    1382 TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTC 1441
      |||
Qy    601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT 660
      |||
Db    1442 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC 1501
      |||
Qy    661 GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720
      |||
Db    1502 GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT 1561
      |||
Qy    721 TCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCCTGCTCCTC 780
      |||
Db    1562 ACCGGCCGCCGAGGTGACCTGGCCACCATTATGCGCATGAACCGGCCCTTCCTGCTTCTC 1621
      |||
Qy    781 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 840
      |||
Db    1622 ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG 1681
      |||
Qy    841 GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC 900
      |||
Db    1682 GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC 1717
      |||
Qy    901 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT 960
      |||
Db    1718 TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC 1777
      |||
Qy    961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020
      |||
Db    1778 GAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTACATTTGGAGCCTG 1837
      |||
Qy    1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTGGGCG 1080
      |||
Db    1838 GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCGGGCGCCTCGGCG 1897
      |||
Qy    1081 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
      |||
Db    1898 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC 1957
      |||
Qy    1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
      |||
Db    1958 AAGCCCAAGGTGGAGGAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 2014

```

RESULT 15

BT007245

LOCUS BT007245 1173 bp mRNA linear PRI 13-MAY-2003

DEFINITION Homo sapiens transforming growth factor, beta 1 (Camurati-Engelmann disease) mRNA, complete cds.

ACCESSION BT007245

VERSION BT007245.1 GI:30583328

KEYWORDS FLI\_CDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1173)

AUTHORS Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,  
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,  
Phelan,M. and Farmer,A.

TITLE Cloning of human full-length CDSs in BD Creator(TM) System Donor  
vector

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1173)

AUTHORS Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,  
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,  
Phelan,M. and Farmer,A.

TITLE Direct Submission

JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow  
Circle, Palo Alto, CA 94303, USA

COMMENT This CDS clone is a part of a collection of human full length  
expression clones generated by BD Biosciences Clontech and the  
Harvard Institute of Proteomics. Each CDS has been cloned in two  
forms: with and without stop-codon (to allow fusion with C-terminal  
tag). The CDS has been directionally cloned using BD In-Fusion(TM)  
cloning system between the SalI and HindIII sites of the pDNR-DUAL  
vector. Additional sequences in the clone: 'ACC' after SalI site  
and before 'ATG' to provide Kozak consensus sequence; 'GG' after  
last codon and before HindIII site to maintain reading frame.  
Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES

source Location/Qualifiers

1. .1173

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="GH00247X1.0"

/clone\_lib="BD Creator(TM) CDS Library derived from MGC  
collection"

/lab\_host="DH5alpha T1 resistant"

/note="Vector: pDNR-Dual"

CDS

1. .1173

/codon\_start=1

/product="transforming growth factor, beta 1  
(Camurati-Engelmann disease)"

/protein\_id="AAP35909.1"

/db\_xref="GI:30583329"

/translation="MPPSGLRLLLLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK  
RIEAIRGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD  
YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRL  
KLKVEQHVELYQKYSNNSWRYLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFR  
LSAHCSCDSRDNTLQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRH  
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT  
QYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS"

BASE COUNT 238 a 380 c 352 g 203 t

ORIGIN

Query Match 79.7%; Score 954.6; DB 9; Length 1173;  
Best Local Similarity 88.5%; Pred. No. 3.1e-165;  
Matches 1057; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

Qy	1	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGGCTGCTGTGGCTGCTAGTG	60
Db	1	ATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTACTGGTG	60
Qy	61	CTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	120
Db	61	CTGACGCCTGGCCGGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	120
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	121	GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	180
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	181	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG	240
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCCGAGCCAGAG	300
Db	241	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCCGAGCCTGAG	300
Qy	301	GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	301	GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC	360
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTC	420
Db	361	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	420
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	421	CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	480
Qy	481	AAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	540
Db	481	AAGTTAAAAGTGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	540
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	600
Db	541	TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTGATGTC	600
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	601	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	660
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	661	GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT	720
Qy	721	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTTCCTGCTCCTC	780
Db	721	ACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTTCCTGCTTCTC	780
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	840
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900





2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:\*  
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1135.8	94.9	1326	24	AAD22696	Porcine transformi
2	958.2	80.1	2742	22	AAI58342	Human polynucleoti
3	956.6	79.9	1559	13	AAQ20289	Sequence encoding
4	956.6	79.9	1561	11	AAQ03268	Simian transformin
5	956.6	79.9	1821	12	AAQ13392	Human pro-TGF-beta
6	956.6	79.9	2537	7	AAN60972	Sequence encoding
7	956.6	79.9	2537	11	AAQ03301	cDNA encoding huma
8	956.6	79.9	2537	11	AAQ02814	Sequence of pre-TG
9	956.6	79.9	2537	17	AAT15720	Pre-transforming g
10	955	79.8	2527	25	ABQ76674	Androgen receptor
11	955	79.8	2537	15	AAQ56923	Human pre-TGF-beta
12	953.4	79.6	1560	9	AAN81084	Coding sequence of
13	953.4	79.6	1560	11	AAQ03508	Simian Transformin
14	951.8	79.5	2537	19	AAV52933	Human pre-transfor
15	945.2	79.0	1176	25	ABV75391	TGFB1 Arg25Pro pol
16	943.6	78.8	1176	24	ABZ35738	Human TGF beta 1 p
17	943.6	78.8	1176	24	ABX09981	Human TGFbeta1 DNA
18	943.6	78.8	1176	24	ABV78162	Human TGF beta 1 D
19	943.6	78.8	1176	24	ABL91703	Human polynucleoti
20	943.6	78.8	1176	25	ABV75392	TGFB1 Arg25Pro pol
21	943.6	78.8	1303	11	AAQ09317	Monkey transformin
22	943.6	78.8	2745	16	AAT05876	cDNA encoding tran
23	943.6	78.8	2745	22	AAH28216	Nucleotide sequenc

24	943.6	78.8	4105	15	AAQ55624	TGFbeta1 5'-UTR-CD
25	940.4	78.6	1571	11	AAQ03269	Human transforming
26	938.8	78.4	1569	9	AAN81085	Coding sequence of
27	937.2	78.3	1569	11	AAQ03509	Human Transforming
28	927.8	77.5	1561	11	AAQ04908	Sequence encoding
29	924.6	77.2	1561	13	AAQ29177	TGF-beta 1/beta 2
30	886.4	74.1	1565	13	AAQ29178	TGF-beta 1. Homo
31	804.6	67.2	2208	13	AAQ20291	Sequence encoding
32	803	67.1	2206	11	AAQ03510	Human Transforming
33	803	67.1	2207	11	AAQ03511	Hybrid transformin
34	801.4	67.0	2207	11	AAQ05127	Human TGF-Beta1/TG
35	790	66.0	2217	10	AAN90768	Sequence of human
36	783.8	65.5	2200	16	AAT04115	Simian-human hybri
c 37	783	65.4	2765	22	AAI60128	Human polynucleoti
c 38	683.2	57.1	2773	23	AAS84421	DNA encoding novel
39	678.8	56.7	834	12	AAQ12192	Sequence encoding
40	654	54.6	1376	24	ABK90341	DNA encoding LAP-m
41	652.4	54.5	1389	24	ABK90344	DNA encoding LAP-h
42	588.2	49.1	1352	24	ABK90342	DNA encoding mIFNB
43	586.6	49.0	1350	24	ABK90343	DNA encoding huIFN
44	524.6	43.8	875	23	AAS70979	DNA encoding novel
45	342.4	28.6	489	24	ABL99528	Target canine gene

#### ALIGNMENTS

##### RESULT 1

AAD22696

ID AAD22696 standard; cDNA; 1326 BP.

XX

AC AAD22696;

XX

DT 26-FEB-2002 (first entry)

XX

DE Porcine transforming growth factor beta 1 (TGF-beta1) cDNA.

XX

KW Porcine; transforming growth factor beta 1; TGF-beta1; gene therapy;  
 KW IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;  
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
 KW diabetes mellitus; sarcoidosis; psoriasis; dermatological; ss.

XX

OS Sus scrofa.

XX

FH Key Location/Qualifiers

FT CDS 16..1188

FT /\*tag= a

FT /product= "Porcine TGF-beta1 mutant protein"

XX

PN WO200181404-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US12980.

XX

PR 20-APR-2000; 2000US-199014P.

XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Strober W, Nakamura K, Kitani A, Fuss IJ;

XX

DR WPI; 2002-026155/03.

DR P-PSDB; AAE13596.

XX

PT Composition for treating autoimmune diseases e.g. inflammatory bowel

PT disease in humans, comprises vector containing transforming growth

PT factor-beta under the control of inducible promoter -

XX

PS Claim 1; Fig 1; 78pp; English.

XX

CC The invention relates to a composition containing a vector comprising a  
CC gene encoding a regulatory transcription factor under the control of a  
CC promoter encoding a transforming growth factor-beta (TGF-beta). The  
CC vector is useful for expressing TGF-beta, such as TGF-beta1, TGF-beta2  
CC or TGF-beta3, its variants or homologues, by transfecting a cell which  
CC is part of a host suspected of having an autoimmune disease, especially  
CC inflammatory bowel disease (IBD), under conditions such that the  
CC polypeptide encoded by the nucleic acid sequence in the vector is  
CC expressed. The vector is delivered using a delivery system. The delivery  
CC of the vector results in substantial elimination of symptoms of the  
CC autoimmune disease and increased production of IL-10 by the host. The  
CC composition is useful for treating various diseases with an autoimmune  
CC component such as multiple sclerosis, rheumatoid arthritis, systemic  
CC lupus erythematosus, insulin-dependent diabetes mellitus, sarcoidosis  
CC and psoriasis, and also for assaying the expression of a gene in a cell.  
CC The vector is further useful for screening of the effect of test  
CC compounds on cytokine (e.g. TGF-beta) expression of transfected cells.  
CC The present sequence is a cDNA encoding porcine TGF-beta1 mutant.

XX

SQ Sequence 1326 BP; 263 A; 438 C; 392 G; 233 T; 0 other;

Query Match 94.9%; Score 1135.8; DB 24; Length 1326;  
Best Local Similarity 97.8%; Pred. No. 4.6e-231;  
Matches 1171; Conservative 0; Mismatches 2; Indels 24; Gaps 1;

Qy	1	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	60
Db	16	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	75
Qy	61	CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	120
Db	76	CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	135
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	136	GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	195
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	196	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	255
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG	300
Db	256	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG	315

Qy	301	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	316	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	375
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTC	420
Db	376	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTC	435
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	436	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	495
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	540
Db	496	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	555
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC	600
Db	556	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC	615
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	616	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	675
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAT	720
Db	676	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAT	735
Qy	721	TCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	780
Db	736	TCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	795
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	796	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	855
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	856	GATACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC	891
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	960
Db	892	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	951
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	952	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1011
Qy	1021	GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCCGGCG	1080
Db	1012	GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCCGGCG	1071
Qy	1081	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1140
Db	1072	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1131

Qy 1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197  
 |||  
 Db 1132 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1188

RESULT 2

AAI58342

ID AAI58342 standard; cDNA; 2742 BP.

XX

AC AAI58342;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 545.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

PF 26-DEC-2000; 2000WO-US34263.

XX

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI; 2001-442253/47.

DR P-PSDB; AAM39186.

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -

XX

PS Claim 1; SEQ ID NO 545; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

XX

SQ Sequence 2742 BP; 526 A; 938 C; 799 G; 479 T; 0 other;

Query Match 80.1%; Score 958.2; DB 22; Length 2742;  
Best Local Similarity 88.6%; Pred. No. 2e-193;  
Matches 1060; Conservative 0; Mismatches 113; Indels 24; Gaps 1;

```
Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     842 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 901

Qy      61 CTGACGCCTGGCCGGCCGGCCCGGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     902 CTGACGCCTGGCCGGCCGGCCCGGCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 961

Qy     121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     962 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 1021

Qy     181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1022 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 1081

Qy     241 TACAACAGTACCCGCGACCGGGTAGCCGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1082 TACAACAGCACCCGCGACCGGGTGGCCGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 1141

Qy     301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1142 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC 1201

Qy     361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420
      ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1202 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 1261

Qy     421 CGGGAAGCGGTGCCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
      || ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1262 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 1321

Qy     481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC 540
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1322 AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA 1381

Qy     541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTCACCGGAGTGGCTGTCCTTTGATGTC 600
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

Db      1382 TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTC 1441
Qy      601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT 660
      |||||
Db      1442 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC 1501
Qy      661 GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720
      |||||
Db      1502 GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT 1561
Qy      721 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
      |
Db      1562 ACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTC 1621
Qy      781 ATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 840
      |||||
Db      1622 ATGGCCACCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG 1681
Qy      841 GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC 900
      ||
Db      1682 GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC 1717
Qy      901 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT 960
      |||||
Db      1718 TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC 1777
Qy      961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020
      ||
Db      1778 GAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCCTG 1837
Qy      1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTGGCG 1080
      |||||
Db      1838 GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCG 1897
Qy      1081 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
      |||||
Db      1898 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC 1957
Qy      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
      |||||
Db      1958 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 2014

```

# RESULT 3

AAQ20289

ID AAQ20289 standard; cDNA; 1559 BP.

XX

AC AAQ20289;

XX

DT 25-MAR-2003 (updated)

DT 16-APR-1992 (first entry)

XX

DE Sequence encoding simian transforming growth factor (TGF) beta-1.

XX

KW Hypertension therapy; hypotensive agent; blood pressure modulator;

KW ss.

XX



```

OS   Monkey.
XX
FH   Key          Location/Qualifiers
FT   CDS          262..282
FT               /*tag= a
FT   sig_peptide  283..324
FT               /*tag= b
FT   CDS          325..1098
FT               /*tag= c
FT   mat_peptide  1099..1436
FT               /*tag= d
XX
PN   WO9119513-A.
XX
PD   26-DEC-1991.
XX
PF   20-JUN-1991;  91WO-US04449.
XX
PR   20-JUN-1990;  90US-0541221.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Oleson FB,  Comerkeski CR;
XX
DR   WPI; 1992-024199/03.
DR   P-PSDB; AAR20124.
XX
PT   Use of transforming growth factor (TGF)-beta and their
PT   antagonists - for modulating blood pressure, for treating
PT   hypertension and hypotension
XX
PS   Disclosure; Fig 1; 42pp; English.
XX
CC   A new method for treating hypertension comprises administering a
CC   transforming growth factor (TGF)-beta to an individual at a dose
CC   effective for lowering blood pressure; the TGF-beta may be e.g.
CC   mature TGF-beta, TGF-beta2, a mature TGF-beta1/beta2 hybrid, TGF-
CC   beta1 precursor, a latent TGF-beta2 precursor, hybrid TGF-beta1/TGF-
CC   beta2 precursor, a latent TGF-beta1 complex or a latent TGF-beta2
CC   complex.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1559 BP; 300 A; 546 C; 446 G; 267 T; 0 other;

Query Match          79.9%;  Score 956.6;  DB 13;  Length 1559;
Best Local Similarity 88.5%;  Pred. No. 4e-193;
Matches 1059;  Conservative  0;  Mismatches 114;  Indels  24;  Gaps  1;

Qy          1  ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG  60
   ||| |||| || ||||||||| ||||||||||| ||||||||||| || |||
Db          261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG  320

Qy          61  CTGACGCCTGGCCGGCCGGCCGCCGACTGTCCACCTGCAAGACCATCGACATGGAGCTG  120
   ||||||| ||||||||||| |||| ||||||||||| ||||||||||| |||
Db          321 CTGACGCCTAGCCGGCCGGCCGCAGGACTATCCACCTGCAAGACTATCGACATGGAGCTG  380

Qy          121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC  180

```

Db	381	 GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	440
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	441	 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG	500
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGCCAGAG	300
Db	501	 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGGAGCCGAGCCCGAACCGGAG	560
Qy	301	GCGGACTACTACGCCAAGGAGGTCAACCGCTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	561	 GCCGACTACTACGCCAAGGAGGTCAACCGCTGCTAATGGTGGAAACCCACAACGAAATC	620
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGCGAGCTC	420
Db	621	 TATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	681	 CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC	540
Db	741	 AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA	800
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC	600
Db	801	 TACCTCAGCAACCGGCTGCTGGCGCCCAGCAACTCGCCGAGTGGTTGTCTTTGATGTC	860
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCTCAGT	660
Db	861	 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCTTAGC	920
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	921	 GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT	980
Qy	721	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	780
Db	981	 ACCGGCCGCCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCCTTCCTGCTTCTC	1040
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1041	 ATGGCCACCCCGCTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1100
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGAATAAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1101	 GACACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC	1136
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT	960
Db	1137	 TGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1196
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020

Db	1197	GAGCCCAAGGGCTACCATGCCAACTTCTGCCTGGGGCCCTGTCCCTACATTTGGAGCCTG	1256
Qy	1021	GACACTCAGTACAGCAAGGTCTTGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCG	1080
Db	1257	GACACGCAGTACAGCAAGGTCTTGCCCTGTACAACCAGCATAAACCGGGCGCCTCGGCG	1316
Qy	1081	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1140
Db	1317	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1376
Qy	1141	AAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1197
Db	1377	AAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAATGCAGCTGA	1433

#### RESULT 4

AAQ03268

ID AAQ03268 standard; DNA; 1561 BP.

XX

AC      AAQ03268;

XX

DT 25-MAR-2003 (updated)

DT 12-AUG-1990 (first entry)

XX

DE Simian transforming growth factor-beta cDNA.

XX

KW Transforming growth factor-beta; psoriasis; TGF-beta; ss.

XX

OS Monkey.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	sig peptide	283..324
----	-------------	----------

```
FT          /*tag= a
```

FT	mat_peptide	1096..1431
----	-------------	------------

```
FT /*tag= b
```

FT /product=human transforming growth factor-beta

XX

PN EP353772-A.

XX

PD 07-FEB-1990.

XX

PF 04-AUG-1989; 89EP-0114458.

XX

PR 05-AUG-1988; 88US-0229133.

XX

PA (ONCO ) ONCOGEN LP.

XX

PI Twardzik DR, Purchio AF, Ranchalis JE, Stevens V;

XX

DR WPI; 1990-038499/06.

DR P-PSDB; AAR03743.

XX

PT Inhibition of proliferation of epidermal cells -

PT used to treat psoriasis by contacting cells with compositions

PT containing transforming growth factor-beta.

XX

PS Disclosure; fig 1; 20pp; English.

XX  
CC TGF-beta may be used in the treatment of hyperplasia  
CC associated with acanthosis-categorised skin diseases, and  
CC in alleviating psoriatic symptoms associated with cytokine-  
CC induced phenomena. See also AAQ03269 and AAR03750.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX

SQ Sequence 1561 BP; 301 A; 547 C; 446 G; 267 T; 0 other;

Query Match 79.9%; Score 956.6; DB 11; Length 1561;  
Best Local Similarity 88.5%; Pred. No. 4e-193;  
Matches 1059; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

```
Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| |||| || ||||||||||| ||||||||||| ||||||||||| ||| |||
Db     262 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 321

Qy      61 CTGACGCCTTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||||||||| ||||||||||| ||||| ||||||||||| ||||||||||| |||
Db     322 CTGACGCCTAGCCGGCCGGCCCGCAGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 381

Qy     121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||||||||||| ||||||||||| ||||| ||||||||||| ||||||||||| |||
Db     382 GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 441

Qy     181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||||||||||| ||||||||||| ||||||||||| ||||||||||| ||||| ||| |||
Db     442 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 501

Qy     241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCAGCCAGAG 300
      ||||||||| ||||||||||| ||||||||| ||||| || ||||||||||| ||| |||
Db     502 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCGGAGCCGGAGCCCGAACCGGAG 561

Qy     301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      || ||||||||||| ||||||||||| ||||||||||| ||||||||||| || ||| |||||
Db     562 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC 621

Qy     361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420
      ||||| || ||||| || || ||||| ||||||| ||||||| || |||||
Db     622 TATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 681

Qy     421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
      || ||||| || ||||||||||| ||||||||||| ||||||||||| |||||
Db     682 CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 741

Qy     481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC 540
      ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||||||||||
Db     742 AAGTTAAAAGTGGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA 801

Qy     541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTGATGTC 600
      ||||||||||| ||||||||||| ||||| ||||| ||||||||||| ||||| |||||||
Db     802 TACCTCAGCAACCGGCTGCTGGCGCCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC 861

Qy     601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT 660
      ||||||||||| ||||||||||| ||||| ||||| || ||||| ||||||||| |||
Db     862 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC 921
```

Qy	661	GCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	922	GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGACATCAACGGGTTCACT	981
Qy	721	TCTGGCCGCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTC	780
Db	982	ACCGGCCGCGCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCCTTCTGCTTCTC	1041
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1042	ATGGCCACCCCACTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1101
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1102	GACACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC	1137
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGATTAT	960
Db	1138	TGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGCTGGAAGTGATCCAC	1197
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1198	GAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATTTGGAGCCTG	1257
Qy	1021	GCACTCAGTACAGCAAGGTCTCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCTGGCG	1080
Db	1258	GACACGAGTACAGCAAGGTCTCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCTGGCG	1317
Qy	1081	GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGCGCG	1140
Db	1318	GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGCGCG	1377
Qy	1141	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAAGTGCAGCTGA	1197
Db	1378	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAATGCAGCTGA	1434

## RESULT 5

AAQ13392

ID AAQ13392 standard; DNA; 1821 BP.

XX

AC    AAQ13392;

XX

DT 20-NOV-1991 (first entry)

XX

DE Human pro-TGF-beta 1 gene.

XX

KW Osteogenetic; tumoricidal; ss.

 $\bar{X}\bar{X}$ 

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT CDS 512..1684

```
FT          /*tag=  a
```

FT	sig_peptide	512..598
----	-------------	----------

```
FT                               /*tag=  b
```

FT misc RNA 599..1684

FT /\*tag= c  
FT /note= "pro-TGF-beta 1"  
FT mat\_peptide 1346..1684  
FT /\*tag= e  
FT /note= "TGF-beta 1"

XX  
PN JP03180192-A.

XX  
PD 06-AUG-1991.

XX  
PF 07-DEC-1989; 89JP-0318243.

XX  
PR 07-DEC-1989; 89JP-0318243.

XX  
PA (KIRI ) KIRIN BREWERY KK.

XX  
DR WPI; 1991-271579/37.  
DR P-PSDB; AAR13813.

XX  
PT Human pro-TGF-beta 1 prodn., for osteo-genetic activity - by  
PT preparing DNA chain contg. base sequence coding for human  
PT pre:pro-TGF-beta 1, forming expression vector etc.

XX  
PS Claim 1; Fig 1; 16pp; Japanese.

XX  
CC The DNA sequence encodes human prepro-TGF-beta 1 which can be  
CC produced by recombinant methods, it has osteogenetic and  
CC tumoricidal activity.

XX  
SQ Sequence 1821 BP; 326 A; 679 C; 508 G; 308 T; 0 other;

Query Match 79.9%; Score 956.6; DB 12; Length 1821;  
Best Local Similarity 88.5%; Pred. No. 4.1e-193;  
Matches 1059; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

Qy 1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60  
||| |||| || ||||||||| ||||||||| ||||||||| |||  
Db 512 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 571  
  
Qy 61 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120  
||||||| ||||| ||||||||| ||||| ||||||||| |||||  
Db 572 CTGACGCCTGGCCCGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 631  
  
Qy 121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180  
||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||  
Db 632 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTGGCC 691  
  
Qy 181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240  
||||||| ||||||||| ||||||||| ||||||||| ||||| ||| |||  
Db 692 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 751  
  
Qy 241 TACAACAGTACCCGCGACCCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG 300  
||||||| ||||||||| ||||||||| ||||| ||||||||| |||||  
Db 752 TACAACAGCACCCGCGACCCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 811  
  
Qy 301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACCAAATC 360  
|| ||||||||| ||||||||| ||||||||| ||| ||||| |||||

Db 812 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC 871

Qy 361 TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420  
 ||||| || ||||| | | ||||| ||||| ||||| ||||| ||||| |||||

Db 872 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 931

Qy 421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480  
 || ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||

Db 932 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 991

Qy 481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC 540  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 992 AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA 1051

Qy 541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC 600  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1052 TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTATGATGTC 1111

Qy 601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGT 660  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1112 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTCGCTTAGC 1171

Qy 661 GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAT 720  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1172 GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCAT 1231

Qy 721 TCTGGCCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780  
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1232 ACCGGCCCGCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTC 1291

Qy 781 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 840  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1292 ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG 1351

Qy 841 GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC 900  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1352 GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC 1387

Qy 901 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT 960  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1388 TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC 1447

Qy 961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1448 GAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCTACATTTGGAGCCTG 1507

Qy 1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCG 1080  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1508 GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCG 1567

Qy 1081 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1568 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC 1627

Qy 1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1628 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1684

RESULT 6

AAN60972

ID AAN60972 standard; cDNA; 2537 BP.

XX

AC AAN60972;

XX

DT 31-OCT-2002 (updated)

DT 28-OCT-1991 (first entry)

XX

DE Sequence encoding preTGF-beta.

XX

KW Transforming growth factor beta; cancer; wound healing.

XX

OS Unidentified.

XX

FH Key Location/Qualifiers

FT misc\_structure 37..113

FT /\*tag= a

FT /note= "Sequence can form stable hairpin loops"

FT CDS 842..2014

FT /\*tag= b

FT mat\_peptide 1676..2011

FT /\*tag= c

XX

PN EP200341-A.

XX

PD 10-DEC-1986.

XX

PF 21-MAR-1986; 86EP-0302112.

XX

PR 22-MAR-1985; 85US-0715142.

PR 13-MAR-1987; 87US-0025423.

XX

PA (GETH ) GENENTECH INC.

XX

PI Derynck RMA;

XX

DR WPI; 1986-326875/50.

DR P-PSDB; AAP61468.

XX

PT TGF-beta prodn. from transformed hosts - useful esp. for treating

PT wounds (J6 2/9/86).

XX

PS Disclosure; Fig 1b; 26pp; English.

XX

CC The gene product is known to stimulate cell proliferation and  
 CC inhibit anchorage-dependent growth of a variety of human cancer cell  
 CC lines, it is esp. useful in treatment of burns and the promotion of  
 CC surface and internal wound healing. TGF-beta may be expressed from a  
 CC transformed CHO cell line.

CC (Updated on 31-OCT-2002 to add missing OS field.)

XX

SQ Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match

79.9%; Score 956.6; DB 7; Length 2537;



Best Local Similarity 88.5%; Pred. No. 4.3e-193;  
Matches 1059; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

```
Qy      1 ATGGCGCCTTCGGGGCTGCGGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| |||| || ||||||||||| ||||||||||| ||||||||||| ||| |||
Db     842 ATGCCGCCCTCCGGGCTGCGGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 901

Qy      61 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||
Db     902 CTGACGCCTGGCCCGCCGGCCCGGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 961

Qy     121 GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||
Db     962 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 1021

Qy     181 AGCCCCCAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||||||||||| ||||||||||| ||||||||||| |||||| ||| |||
Db    1022 AGCCCCCAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 1081

Qy     241 TACAACAGTACCCGCGACCCGGGTAGCCGGGGAAAGTGTGCAACCGAGCCCGAGCCAGAG 300
      |||||| ||||||||||| |||||| ||| ||||||||||| |||
Db    1082 TACAACAGCACCCGCGACCCGGGTGGCCGGGGAGAGTGCAGAACCGAGCCCGAGCCTGAG 1141

Qy     301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      || ||||||||||| ||||||||||| |||||| ||| |||
Db    1142 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC 1201

Qy     361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420
      ||||| || |||||| | |||||| ||||||| |||||| ||| |||
Db    1202 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 1261

Qy     421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
      || |||||| || |||||| || |||||| ||||||||||| |||||||||||
Db    1262 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 1321

Qy     481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC 540
      ||||||||||| ||||||||||| |||||| ||| |||||| |||
Db    1322 AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA 1381

Qy     541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC 600
      ||||||||||| ||||||||||| |||||| ||| |||||| |||
Db    1382 TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTC 1441

Qy     601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT 660
      ||||||||||| |||||| ||| ||| || ||| ||| ||| |||
Db    1442 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC 1501

Qy     661 GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720
      ||||| ||| |||||| ||||||||||| || |||| || |||||||
Db    1502 GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT 1561

Qy     721 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
      | |||||| ||||||||||| ||||||||||| |||||| |||
Db    1562 ACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTCCTC 1621

Qy     781 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 840
      ||||||||||| ||||||||||| ||||| ||||||||||| |||||||
```

```

Db      1622 ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG 1681
Qy      841 GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC 900
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1682 GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC 1717
Qy      901 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT 960
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1718 TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC 1777
Qy      961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1778 GAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCCTG 1837
Qy      1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCG 1080
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1838 GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCG 1897
Qy      1081 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1898 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC 1957
Qy      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1958 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 2014

```

# RESULT 7

AAQ03301

ID AAQ03301 standard; DNA; 2537 BP.

XX

AC AAQ03301;

XX

DT 25-MAR-2003 (updated)

DT 05-AUG-1990 (first entry)

XX

DE cDNA encoding human pre-transforming growth factor-beta-1 (pre-TGF-beta-1).

XX

KW Transforming growth factor-beta-1 (TGF-beta-1);

KW neoplastic cell line inhibition;

KW EGF-potentiased anchorage-independent growth;

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	842..2014
----	-----	-----------

FT		/*tag= a
----	--	----------

FT	mat_peptide	1676..2011
----	-------------	------------

FT		/*tag= b
----	--	----------

FT	misc_difference	37..113
----	-----------------	---------

FT		/*tag= c
----	--	----------

FT		/note="stable hairpin loops"
----	--	------------------------------

FT	misc_feature	2015..2100
----	--------------	------------

FT		/*tag= d
----	--	----------

FT		/note="G-C rich sequence
----	--	--------------------------

FT		and a downstream TATA-like sequence"
----	--	--------------------------------------

XX  
 PN US4886747-A.  
 XX  
 PD 12-DEC-1989.  
 XX  
 PF 13-MAR-1987; 87US-0025423.  
 XX  
 PR 13-MAR-1987; 87US-0025423.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Derynck RMA, Goeddel DV;  
 XX  
 DR WPI; 1990-051338/07.  
 DR P-PSDB; AAR05258.  
 XX  
 PT Nucleic acid encoding transforming growth factor-beta -  
 PT cloned into expression vectors for expression in eukaryotic host  
 PT cells for therapeutic use  
 XX  
 PS Disclosure; Fig 1b; 28pp; English.  
 XX  
 CC It was obtained by an analysis of several overlapping cDNAs and gene  
 CC fragments, leading to the detn. of a continuous sequence corresp. to the  
 CC TGF-beta-1 precursor mRNA. It is useful in constructing vectors that  
 CC encode biologically active transforming growth factor (TGF-beta),  
 CC operably linked to DNA that encodes a secretory leader (SL). It, or a  
 CC nucleic acid capable of hybridising with it, can also be labelled and  
 CC used in diagnostic assays for DNA or mRNA encoding TGF-beta or related  
 CC proteins.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match 79.9%; Score 956.6; DB 11; Length 2537;  
 Best Local Similarity 88.5%; Pred. No. 4.3e-193;  
 Matches 1059; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

Qy	1	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	60
Db	842	ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG	901
Qy	61	CTGACGCCTGGCCGCGCCGCGCCGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	120
Db	902	CTGACGCCTGGCCCGCGCGCCGCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	961
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	962	GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCGCCAGATCCTGTCCAAGCTGCGGCTCGCC	1021
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCGCCGCGCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	1022	AGCCCCCGAGCCAGGGGGAGGTGCCGCGCCGCGCCCGCTGCCCGAGGCGGTGCTCGCCCTG	1081
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCCGAGCCAGAG	300
Db	1082	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCCGAGCCTGAG	1141

Qy	301	GCGGACTACTACGCCAAGGAGGTCACCCGCTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	1142	GCCGACTACTACGCCAAGGAGGTCACCCGCTGCTAATGGTGGAAACCCACAACGAAATC	1201
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	420
Db	1202	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	1261
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	1262	CGAGAAGCGGTACCTGAACCCGTGTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	1321
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	540
Db	1322	AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	1381
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC	600
Db	1382	TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTGATGTC	1441
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGT	660
Db	1442	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTCGCCTTAGC	1501
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1502	GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT	1561
Qy	721	TCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	780
Db	1562	ACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTC	1621
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1622	ATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1681
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1682	GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC	1717
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	960
Db	1718	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1777
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1778	GAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCCTG	1837
Qy	1021	GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCG	1080
Db	1838	GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCGGGCGCCTCGGCG	1897
Qy	1081	GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1140
Db	1898	GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC	1957

QY 1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197  
 |||  
 Db 1958 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAAGTGCAGCTGA 2014

RESULT 8

AAQ02814

ID AAQ02814 standard; cDNA; 2537 BP.

XX

AC AAQ02814;

XX

DT 25-MAR-2003 (updated)

DT 31-OCT-2002 (updated)

DT 31-MAY-1989 (first entry)

XX

DE Sequence of pre-TGF-beta1 cDNA.

XX

KW Transforming growth factor beta-3 (TGF beta 3); tumour cells; growth  
 KW inhibition.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 842..2011

FT /\*tag= a

FT /label=pre-TGF beta 1

FT CDS 1677..2011

FT /\*tag= b

FT /label=mature TGF-beta 1

FT GC\_signal 2015..2092

FT /\*tag= c

FT misc\_feature 2093..2099

FT /\*tag= d

FT /label=TATA-like sequence

FT stem\_loop 37..113

FT /\*tag= e

FT misc\_feature 863..911

FT /\*tag= f

FT /label=hydrophobic domain

XX

PN WO8912101-A.

XX

PD 14-DEC-1989.

XX

PF 08-JUN-1988; 88WO-US01945.

XX

PR 08-JUN-1988; 88WO-US01945.

XX

PA (GETH ) GENENTECH INC.

XX

PI Dernyck RMA, Goeddel DV;

XX

DR WPI; 1990-007474/01.

DR P-PSDB; AAR04034.

XX

PT Nucleotide sequence encoding transforming growth factor beta-3 -used as a  
 PT probe, or to produce TGF beta 3, for growth inhibition of certain normal

PT and neoplastic cells, eq A549.

PS Disclosure; Fig. 1b; 61pp; English.

Sequence encodes the 390 amino acid (AA) precursor transforming growth factor-beta 1 (pre-TGF-beta 1) polypeptide. The 5' untranslated region of the TGF-beta 1 mRNA is 841 bases long, is purine rich and has a region of potential secondary structure. The TATA-like sequence in the 3' untranslated region of the gene is presumably a polyadenylation signal. Mature TGF-beta 1 comprises the C-terminal 112 AA's of pre-TGF-beta 1 and is cleaved at the Arg-Arg dipeptide preceding its NH2 terminus. The nucleic acid encoding the second subtype of TGF-beta (TGF-beta 3) is useful as a probe or to produce TGF-beta 3 for inhibition of growth of normal and neoplastic cells.

CC (Updated on 31-OCT-2002 to add missing OS field.)

CC (Updated on 25-MAR-2003 to correct PR field.)

CC (Updated on 25-MAR-2003 to correct PI field.)

SQ Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match 79.9%; Score 956.6; DB 11; Length 2537;  
Best Local Similarity 88.5%; Pred. No. 4.3e-193;  
Matches 1059; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

Qy	1	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	60
Db	842	ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGTGCTGTGGCTACTGGTG	901
Qy	61	CTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	120
Db	902	CTGACGCCTGGCCCCGCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	961
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	962	GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	1021
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	1022	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCGCCCTG	1081
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCAGCCAGAG	300
Db	1082	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAAAACCGGAGCCCAGCCTGAG	1141
Qy	301	GCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	1142	GCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC	1201
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGCGAGCTC	420
Db	1202	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	1261
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	1262	CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	1321
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTTCCTGGCGC	540

Db	1322		AAGTTAAAGTGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA	1381
Qy	541		TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTGATGTC	600
Db	1382		TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTGATGTC	1441
Qy	601		ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	1442		ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	1501
Qy	661		GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1502		GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT	1561
Qy	721		TCTGGCCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	780
Db	1562		ACCGGCCCGCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTC	1621
Qy	781		ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1622		ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1681
Qy	841		GATACCAACTACTGCTTCAGCTCCACGGA CTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1682		GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC	1717
Qy	901		TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT	960
Db	1718		TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1777
Qy	961		GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1778		GAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCCTG	1837
Qy	1021		GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCG	1080
Db	1838		GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCG	1897
Qy	1081		GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1140
Db	1898		GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC	1957
Qy	1141		AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1197
Db	1958		AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	2014

RESULT 9

AAT15720

ID AAT15720 standard; cDNA; 2537 BP.

XX

AC AAT15720;

XX

DT 25-MAR-2003 (updated)

DT 24-JUL-1997 (revised)

DT 25-JAN-1980 (first entry)

```

XX
DE   Pre-transforming growth factor beta 1 cDNA.
XX
KW   transforming growth factor beta 1; wound healing;
KW   recombinant production; ss.
XX
OS   Homo sapiens.
XX
FH   Key          Location/Qualifiers
FT   5'UTR        1..841
FT               /*tag= a
FT   misc_feature  37..113
FT               /*tag= b
FT               /note= "GC-rich region forms stable hairpin loops;
FT                   similar to structural organisation of c-myc RNA,
FT                   could play role in mRNA stability or in
FT                   regulation of transcription"
FT   CDS           842..2014
FT               /*tag= c
FT               /product= pre-TGF_beta_1
FT   mat_peptide   1676..2011
FT               /*tag= d
FT               /product= mature_TGF_beta_1
FT   repeat_region 2015..2100
FT               /*tag= e
FT               /note= "GC-rich region; possibly responsible for the
FT                   fact 3'UTR of mRNA could not be cloned as cDNA;
FT                   may be important for transcription efficiency"
FT   repeat_unit   2019..2023
FT               /*tag= f
FT   TATA_signal   2094..2100
FT               /*tag= g
FT               /note= "TATA-like sequence; no evidence that this
FT                   functions a promoter"
FT   polyA_signal  2514..2520
FT               /*tag= h
FT   misc_signal   2529..2536
FT               /*tag= i
FT               /note= "consensus sequence immediately precedes
FT                   polyA-tail (Benoist et al)"
XX
PN   US5482851-A.
XX
PD   09-JAN-1996.
XX
PF   05-NOV-1993; 93US-0147364.
XX
PR   13-MAR-1987; 87US-0025423.
PR   22-MAR-1985; 85US-0715142.
PR   04-AUG-1989; 89US-0389929.
PR   04-MAR-1992; 92US-0845893.
PR   05-NOV-1993; 93US-0147364.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Derynck RMA, Goeddel DV;
XX

```



DR WPI; 1996-076891/08.  
 DR P-PSDB; AAR90827.  
 XX  
 PT New recombinant human transforming growth factor-beta prods. - produced  
 PT using Chinese hamster ovary cells, for use in diagnostic applications  
 PT or in therapy  
 XX  
 PS Example 3; Fig 1; 26pp; English.  
 XX  
 CC The cDNA encodes the pre-transforming growth factor (TGF) beta 1 protein.  
 CC The nucleotide sequence was obtd. by an analysis of several overlapping  
 CC cDNAs and gene fragments. The DNA is useful for the recombinant  
 CC production of TGF beta 1, which can be used in, e.g. wound healing.  
 CC (Revised entry submitted to correct sequence analysis breakdown.)  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match 79.9%; Score 956.6; DB 17; Length 2537;  
 Best Local Similarity 88.5%; Pred. No. 4.3e-193;  
 Matches 1059; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

```

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      842 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 901

Qy      61 CTGACGCCTGGCCGCGCCGCGCCGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      902 CTGACGCCTGGCCGCGCCGCGCGGCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 961

Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      962 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 1021

Qy      181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCCGCCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1022 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCCGCCCCGCTGCCCGAGGCCGTGCTCGCCCTG 1081

Qy      241 TACAACAGTACCCGCGACCGGGTAGCCGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1082 TACAACAGCACCCGCGACCGGGTGGCCGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 1141

Qy      301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1142 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC 1201

Qy      361 TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420
      ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1202 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 1261

Qy      421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
      || ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1262 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 1321

Qy      481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAATACAGCAATGATTCCTGGCGC 540
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1322 AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA 1381

```

Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC	600
Db	1382	TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTC	1441
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	1442	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTCGCCTTAGC	1501
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAT	720
Db	1502	GCCCACTGTCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT	1561
Qy	721	TCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTC	780
Db	1562	ACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGCTTCTC	1621
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1622	ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1681
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1682	GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC	1717
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT	960
Db	1718	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1777
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1778	GAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCCTG	1837
Qy	1021	GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCCGGCG	1080
Db	1838	GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCCGGCG	1897
Qy	1081	GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1140
Db	1898	GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC	1957
Qy	1141	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1197
Db	1958	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	2014

RESULT 10

ABQ76674

ID ABQ76674 standard; DNA; 2527 BP.

XX

AC ABQ76674;

XX

DT 26-MAR-2003 (first entry)

XX

DE Androgen receptor signalling pathway-associated DNA E00973.

XX

KW Androgen receptor; transactivation; modulator; Smad3; Smad4; Akt; TGF-B;

KW signal transduction pathway; transforming growth factor-B; phosphatase;  
KW tensin; cytostatic; antiproliferative; cellular proliferation; cancer;  
KW E00973; ds.

XX

OS Synthetic.

XX

PN WO200282081-A2.

XX

PD 17-OCT-2002.

XX

PF 05-APR-2002; 2002WO-US11086.

XX

PR 06-APR-2001; 2001US-282266P.

PR 13-MAR-2002; 2002US-365060P.

XX

PA (UYRP ) UNIV ROCHESTER.

XX

PI Chang C;

XX

DR WPI; 2003-046871/04.

XX

PT Modulating androgen receptor activity, by administering a compound that  
PT modulates receptor activity, inhibits receptor-signal transduction  
PT pathway/receptor-coactivator interaction or changes amount or receptor  
PT -

XX

PS Disclosure; Page 225-226; 302pp; English.

XX

CC This invention describes a novel method for modulating androgen receptor  
CC activity or androgen receptor-mediated transactivation activity in a  
CC cell. The method involves administering a compound which causes  
CC modulation of the androgen receptors activity and the inhibition of  
CC interaction between the receptor and a protein involved in a signal  
CC transduction pathway. The compound also inhibits the interaction between  
CC the androgen receptor and a protein selected from Smad3, Smad4, Akt,  
CC transforming growth factor (TGF)-B and phosphatase and tensin homologues  
CC deleted on chromosome 10 (PTEN) or their fragments. The compounds of the  
CC invention have cytostatic and antiproliferative activity. The obtained  
CC composition is useful for treating any disease, where uncontrolled  
CC proliferation or cellular proliferation occurs such as cancer, e.g.  
CC prostate cancer. This sequence represents the androgen receptor  
CC transactivation signalling pathway modulator E00973 described in  
CC the method of the invention.

XX

SQ Sequence 2527 BP; 472 A; 888 C; 735 G; 432 T; 0 other;

Query Match 79.8%; Score 955; DB 25; Length 2527;  
Best Local Similarity 88.4%; Pred. No. 9.3e-193;  
Matches 1058; Conservative 0; Mismatches 115; Indels 24; Gaps 1;

Qy 1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60  
||| |||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 842 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 901  
  
Qy 61 CTGACGCCTGGCCGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 902 CTGACGCCTGGCCGGCCGGCCGGCCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 961

Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTCCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	962	GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	1021
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	1022	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCGAGGCCGTGCTCGCCCTC	1081
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGAAACCGAGCCCGAGCCAGAG	300
Db	1082	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGAGAAACCGAGCCCGAGCCTGAG	1141
Qy	301	GCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGAAAGCGGCAACCAAATC	360
Db	1142	GCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGAAACCCACAACGAAATC	1201
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGCGAGCTC	420
Db	1202	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	1261
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	1262	CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	1321
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC	540
Db	1322	AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA	1381
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTTGATGTC	600
Db	1382	TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTC	1441
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGGCTATAGAGGGTTTTTCGCTCAGT	660
Db	1442	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCTTAGC	1501
Qy	661	GCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1502	GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT	1561
Qy	721	TCTGGCCGCCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCTGCTCCTC	780
Db	1562	ACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGCTTCTC	1621
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1622	ATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1681
Qy	841	GATACCAACTACTGCTTCAGCTCCACGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1682	GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC	1717
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT	960
Db	1718	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1777

```

Qy      961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020
      || |||||
Db      1778 GAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTACATTTGGAGCCTG 1837

Qy      1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCG 1080
      |||||
Db      1838 GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCGGGCGCCTCGGCG 1897

Qy      1081 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGCCGC 1140
      |||||
Db      1898 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTAACGTGGGCCGC 1957

Qy      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
      |||||
Db      1958 AAGCCCAAGGTGGAGGAGCTGTCCAACATGATCGTGCGCTCCTGCAAGTGCAGCTGA 2014

```

RESULT 11

AAQ56923

ID AAQ56923 standard; cDNA; 2537 BP.

XX

AC AAQ56923;

XX

DT 25-MAR-2003 (updated)

DT 09-JUL-1994 (first entry)

XX

DE Human pre-TGF-beta-1.

XX

KW TGF-beta-1; TGF-beta-2; transforming growth factor beta-1;

KW transforming growth factor beta-3; recombinant; wound healing;

KW vulnerary; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT misc\_structure 47..113

FT /\*tag= a

FT /note= "possible hairpin loop region"

FT CDS 842..2014

FT /\*tag= b

FT mat\_peptide 1676..2011

FT /\*tag= c

FT polyA\_signal 2515..2521

FT /\*tag= d

XX

PN US5284763-A.

XX

PD 08-FEB-1994.

XX

PF 04-MAR-1992; 92US-0845893.

XX

PR 22-MAR-1985; 85US-0715142.

PR 13-MAR-1987; 87US-0025423.

PR 04-AUG-1989; 89US-0389929.

PR 04-MAR-1992; 92US-0845893.

XX

PA (GETH ) GENENTECH INC.

```

XX
PI   Derynk RMA,  Goeddel DV;
XX
DR   WPI; 1994-056343/07.
DR   P-PSDB; AAR46227.
XX
PT   Nucleic acid sequences encoding transforming growth factor-beta -
PT   diagnostic probes, and for use in therapeutics
XX
PS   Disclosure; Fig 1b; 25pp; English.
XX
CC   cDNA sequences were determined for human pre-TGF-beta-1 (AAQ56923),
CC   pig TGF-beta-3 (AAQ56925) and human TGF-beta-3 (AAQ56926), and the
CC   corresponding amino acid sequences were determined (AAR46227-29,
CC   respectively). A genomic fragment corresponding to a human TGF-
CC   beta-1 exon (AAQ56924) was also isolated and its amino acid sequence
CC   determined (AAR46230). The sequences have been used in the
CC   construction of vectors for the expression of recombinant TGF-
CC   beta.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 2537 BP; 473 A; 890 C; 742 G; 432 T; 0 other;

Query Match          79.8%;  Score 955;  DB 15;  Length 2537;
Best Local Similarity 88.4%;  Pred. No. 9.3e-193;
Matches 1058;  Conservative 0;  Mismatches 115;  Indels 24;  Gaps 1;

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| |||| || ||||||||||| ||||||||||| ||||||||||| |||
Db      842 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 901

Qy      61 CTGACGCCTGGCCGCGCCGCGCCGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||||||||||| ||||||| ||||| ||||||||||| ||||||||||| |||
Db      902 CTGACGCCTGGCCGCGCCGCGCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 961

Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||||||||||| ||||||| ||||||| ||||||| ||||||| ||||| |||
Db      962 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 1021

Qy      181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||||||||||| ||||||| ||||||| ||||||| ||||||| ||||| |||
Db      1022 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 1081

Qy      241 TACAACAGTACCCGCGACCGGGTAGCCGGGAAAGTGTGCAACCGAGCCCAGCCAGAG 300
      ||||||| ||||||| ||||||| ||||||| ||||| ||||| ||||||| |||
Db      1082 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACGGAGCCCAGCCTGAG 1141

Qy      301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      || ||||||||||| ||||||| ||||||| ||||||| ||||| ||||| |||
Db      1142 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC 1201

Qy      361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420
      ||||| || ||||||| ||| ||||||| ||||||| ||||||| || |||||||
Db      1202 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 1261

Qy      421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
      || ||||||| || ||||||| || ||||||| ||||||| ||||||| |||

```



```

DT      25-MAR-2003   (updated)
DT      09-OCT-1990   (first entry)
XX
DE      Coding sequence of simian transforming growth factor-beta 1.
XX
KW      Transforming growth factor-beta 1; tumour treatment; ss cDNA.
XX
OS      Cercopithecus aethiops.
XX
FH      Key           Location/Qualifiers
FT      CDS           261..1433
FT                        /*tag= a
FT      sig_peptide   282..323
FT                        /*tag= b
FT      mat_peptide   1095..1433
FT                        /*tag= c
XX
PN      EP293785-A.
XX
PD      07-DEC-1988.
XX
PF      27-MAY-1988;   88EP-0108528.
XX
PR      29-MAY-1987;   87US-0055662.
PR      25-JAN-1988;   88US-0147842.
XX
PA      (ONCO ) ONCOGEN.
PA      (BRIM ) BRISTOL-MYERS CO.
XX
PI      Purchio AG, Gentry L, Twardzik D;
XX
DR      WPI; 1988-347488/49.
DR      P-PSDB; AAP80647.
XX
PT      Prodn. of simian transforming growth factor beta-1 - by culturing
PT      transfected eucaryotic cells, and new precursor proteins, useful for
PT      treating tumours.
XX
PS      Disclosure; Page ?; pp; English.
XX
CC      The cDNA is prepd. from African green monkey cell line BSC-40 and is
CC      expressed in eukaryotic cells in plasmid pSV2. There is 100% homology
CC      between mature simian and human TGF-beta 1. The plasmid also contains
CC      the SV40 promoter and a selection marker, esp. DHFR.
CC      (Updated on 25-MAR-2003 to correct PA field.)
CC      (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ      Sequence 1560 BP; 301 A; 547 C; 445 G; 267 T; 0 other;

Query Match          79.6%;   Score 953.4;   DB 9;   Length 1560;
Best Local Similarity 88.3%;   Pred. No. 1.9e-192;
Matches 1057;   Conservative    0;   Mismatches 116;   Indels    24;   Gaps    1;

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      261 ATGGCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 320

```



Qy	61	CTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	120
Db	321	CTGACGCCTAGCCGGCCGGCCGAGGACTATCCACCTGCAAGACTATCGACATCGAGCTG	380
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	381	GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	440
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	441	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG	500
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG	300
Db	501	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGGAGCCGAGCCCGAACCGGAG	560
Qy	301	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	561	GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC	620
Qy	361	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	420
Db	621	TATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	681	CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	540
Db	741	AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	800
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	600
Db	801	TACCTCAGCAACCGGCTGCTGGCGCCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	860
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGT	660
Db	861	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCTTAGC	920
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	921	GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT	980
Qy	721	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTC	780
Db	981	ACCGGCCGCCGAGGTGACCTGGCCACAATTATGGCATGAACCGGCCCTTCTGCTTCTC	1040
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1041	ATGGCCACCCCGCTGGAGAGGGCCCCAACATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1100
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1101	GACACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC	1136
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	960

Db	1137	TGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1196
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1197	GAGCCCAAGGGCTACCATGCCAACTTCTGCCTGGGGCCCTGTCCCTACATTTGGAGCCTG	1256
Qy	1021	GACACTCAGTACAGCAAGGTCTGTGCTGTGTACAACCAGCACAAACCCGGGCGCGTTCGGCG	1080
Db	1257	GACACGCAGTACAGCAAGGTCTGTGCTGTGTACAACCAGCATAACCCGGGCGCCTTCGGCG	1316
Qy	1081	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1140
Db	1317	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1376
Qy	1141	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1197
Db	1377	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGAAAATGCAGCTGA	1433

RESULT 13

AAQ03508

ID    AAQ03508 standard; DNA; 1560 BP.

XX

AC    AAQ03508;

XX

DT 25-MAR-2003 (updated)

DT 09-JAN-2003 (updated)

DT 14-AUG-1990 (first entry)

XX

DE Simian Transforming growth factor - Betal.

XX

KW HIV; AIDS; SIV; vaccine; AZT; CD4; cytokines; growth

KW factors; ds.

XX

OS    Cebus apella.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT CDS 267..1437

FT /\*tag= a

FT mat peptide 1103..1437

FT /\*tag= b

XX

PN EP356935-A.

XX

PD 07-MAR-1990.

XX

PF 25-AUG-1989; 89EP-0115719.

XX

PR 25-AUG-1988; 88US-0236698.

XX

PA (ONCO ) ONCOGEN LP.

XX

PI Brankovan V, Lioubin M, Purchio A;

XX

DR WPI; 1990-068723/10.

DR P-PSDB; AAR05663.

```

XX   Compsns. contg. transforming growth factor beta -
PT   used for inhibitions of HIV infection and replication in vivo.
XX
PS   Disclosure; Fig 1; 20pp; English.
XX
CC   TGF-beta may be used in vivo to prevent formation of syncytia and
CC   inhibit HIV infection. TGF may also be used with other HIV treatments
CC   (AZT, soluble CD4 etc.).
CC   (Updated on 09-JAN-2003 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1560 BP; 301 A; 547 C; 445 G; 267 T; 0 other;

Query Match          79.6%;  Score 953.4;  DB 11;  Length 1560;
Best Local Similarity 88.3%;  Pred. No. 1.9e-192;
Matches 1057;  Conservative    0;  Mismatches 116;  Indels    24;  Gaps    1;

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG  60
      ||| |||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGTGCTGTGGCTACTGGTG  320

Qy      61 CTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG  120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      321 CTGACGCCTAGCCGGCCGGCCGGCAGGACTATCCACCTGCAAGACTATCGACATCGAGCTG  380

Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC  180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      381 GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC  440

Qy      181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCCTGAGGCAGTACTGGCTCTT  240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      441 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG  500

Qy      241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG  300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      501 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCGGAGCCCGGAGCCCGAACCGGAG  560

Qy      301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGAAAGCGGCAACCAAATC  360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      561 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGAAACCCACAACGAAATC  620

Qy      361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC  420
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db      621 TATGACAAGTTCAAGCAGAGCACACAGCATATATATGTTCTTCAACACATCAGAGCTC  680

Qy      421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC  480
      || ||||| || || ||||| || ||||| || ||||| || ||||| || |||||
Db      681 CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGAGGCTC  740

Qy      481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC  540
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      741 AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA  800

Qy      541 TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC  600
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      801 TACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGAGTGGTTGTCTTTGATGTC  860

```

Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	861	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	920
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	921	GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGACATCAACGGGTTCACT	980
Qy	721	TCTGGCCGCGCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCTGCTCCTC	780
Db	981	ACCGGCCGCGCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCCTTCTGCTTCTC	1040
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1041	ATGGCCACCCCGCTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1100
Qy	841	GATACCAACTACTGCTTCAGCTCCACGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1101	GACACCAACTACTGCTTCAGCTCCAC-----GGAGAAGAAC	1136
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	960
Db	1137	TGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1196
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1197	GAGCCCAAGGGCTACCATGCCAACTTCTGCCTGGGGCCCTGTCCCTACATTTGGAGCCTG	1256
Qy	1021	GCACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTTCGGCG	1080
Db	1257	GACACGAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCGGGCGCCTTCGGCG	1316
Qy	1081	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1140
Db	1317	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1376
Qy	1141	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1197
Db	1377	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGAAAATGCAGCTGA	1433

AAV52933

XX

XX

DT 21-DEC-1998 (first entry)

DE Human pre-transforming growth factor-beta 1 cDNA.

KW Transforming growth factor-beta 1; TGF-beta 1; human; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	842..2014
FT		/*tag= a
FT	mat_peptide	1676..2011
FT		/*tag= b
FT	stem_loop	37..113
FT		/*tag= b
FT		/note= "putative stable hairpin loop"
FT	misc_feature	2015..2100
FT		/*tag= c
FT		/note= "GC-rich sequence"
FT	polyA_signal	2514..2520
FT		/*tag= d
XX		
PN	US5801231-A.	
XX		
PD	01-SEP-1998.	
XX		
PF	30-MAY-1995;	95US-0454468.
XX		
PR	13-MAR-1987;	87US-0025423.
PR	22-MAR-1985;	85US-0715142.
PR	04-AUG-1989;	89US-0389929.
PR	04-MAR-1992;	92US-0845893.
PR	05-NOV-1993;	93US-0147364.
PR	30-MAY-1995;	95US-0454468.
XX		
PA	(GETH ) GENENTECH INC.	
XX		
PI	Derynck RMA, Goeddel DV;	
XX		
DR	WPI; 1998-494840/42.	
DR	P-PSDB; AAW78785.	
XX		
PT	DNA encoding transforming growth factor-beta precursor sequence -	
PT	useful for analysis to perform manipulations to increase yield of	
PT	recombinant production of the protein	
XX		
PS	Example 3; Fig 1B 1-3; 26pp; English.	
XX		
CC	This nucleotide sequence codes for the human transforming growth	
CC	factor-beta 1 precursor (preTGF-beta 1, see AAW78785). It is a	
CC	composite of overlapping cDNA clones isolated from different cDNA	
CC	libraries (placenta, A172 glioblastoma, HT1080 fibroblastoma) using	
CC	TGF-beta exon (see AAV52936) restriction fragments as probes.	
CC	The 3' region of the sequence was determined using cloned genomic	
CC	DNA. The invention relates to the recombinant production of	
CC	TGF-beta. Biologically active TGF-beta is defined as being capable	
CC	of inducing EGF-potentiated anchorage independent growth of target	
CC	cell lines and/or growth inhibition of neoplastic cell lines.	
CC	Nucleic acids encoding TGF-beta have been isolated and cloned into	
CC	vectors which are replicated in bacteria and expressed in	
CC	eukaryotic cells. TGF-beta recovered from transformed cells is	
CC	used in known therapeutic applications. TGF-beta nucleic acids are	
CC	also useful in diagnosis and identification of TGF-beta clones.	
CC	(Updated on 25-MAR-2003 to correct PF field.)	
XX		

SQ Sequence 2537 BP; 475 A; 895 C; 736 G; 431 T; 0 other;

Query Match 79.5%; Score 951.8; DB 19; Length 2537;  
Best Local Similarity 88.2%; Pred. No. 4.4e-192;  
Matches 1056; Conservative 0; Mismatches 117; Indels 24; Gaps 1;

```
Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     842 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 901

Qy      61 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     902 CTGACGCCTGGCCCGCCGGCCCCGGGACTATCCACCTGCAAGACTATCGACATGGAGCAG 961

Qy     121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     962 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 1021

Qy     181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1022 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 1081

Qy     241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGCCAGAG 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1082 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 1141

Qy     301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACCAAATC 360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1142 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACGAAATC 1201

Qy     361 TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGCGAGCTC 420
      ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1202 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 1261

Qy     421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1262 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 1321

Qy     481 AAGTTAAAAAGTGGAGCAGCACGTGGAGCTATAACCAGAAATACAGCAATGATTCTGGCGC 540
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1322 AAGTTAAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA 1381

Qy     541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCCGGAGTGGCTGTCTTTGATGTC 600
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1382 TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTGATGTC 1441

Qy     601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGT 660
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1442 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGAGAAATTGAGGGCTTTCGCCTTAGC 1501

Qy     661 GCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1502 GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT 1561

Qy     721 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    1562 ACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTC 1621
```

```

Qy      781 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 840
      |||
Db      1622 ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG 1681

Qy      841 GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC 900
      || |||
Db      1682 GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC 1717

Qy      901 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT 960
      |||
Db      1718 TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC 1777

Qy      961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020
      || |||
Db      1778 GAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCCTG 1837

Qy      1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCG 1080
      |||
Db      1838 GACACGAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCG 1897

Qy      1081 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
      |||
Db      1898 GCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC 1957

Qy      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
      |||
Db      1958 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 2014

```

RESULT 15

ABV75391

ID ABV75391 standard; DNA; 1176 BP.

XX

AC ABV75391;

XX

DT 18-MAR-2003 (first entry)

XX

DE TGFB1 Arg25Pro polymorphism G-allele nucleotide sequence.

XX

KW Polymorphism; transforming growth factor beta 1; TGFB1; TGFbeta-1;  
renal failure; nephrotropic; human; gene; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..1176

FT /\*tag= a

XX

PN WO200290585-A2.

XX

PD 14-NOV-2002.

XX

PF 08-MAY-2002; 2002WO-GB02066.

XX

PR 09-MAY-2001; 2001GB-0011277.

XX

PA (UYSH-) UNIV SHEFFIELD HALLAM.  
 XX  
 PI El-Nahas AM, Blakemore A, Khalil MS;  
 XX  
 DR WPI; 2003-120560/11.  
 DR P-PSDB; ABB82780.  
 XX  
 PT Determining an individual's susceptibility to the progression of renal  
 PT failure comprises detecting the presence of a genetic polymorphism  
 PT pattern in transforming growth factor beta 1 (TGFB1) gene in a sample  
 PT from the individual -  
 XX  
 PS Claim 49; Page 58; 62pp; English.  
 XX  
 CC The invention relates to determining an individual's susceptibility to  
 CC renal failure and involves detecting the presence of a genetic  
 CC polymorphism pattern in transforming growth factor beta 1 (TGFB1) gene in  
 CC a sample from an individual, where polymorphism pattern is associated  
 CC with renal failure. The method is useful for determining an individual's  
 CC susceptibility to the progression of renal failure. The nucleic acid  
 CC comprising a T(-509)C polymorphism of TGFB1 gene, or a polypeptide  
 CC comprising a sequence of 391 amino acids is useful for preparing a  
 CC medicament for retarding or preventing the progression of renal disease,  
 CC and for drug research purposes for retarding or preventing the  
 CC progression of renal disease. Sequences ABV75386-88 represents the  
 CC nucleotide sequence for the TGFB1 G-allele of the Arg25Pro polymorphism  
 CC of exon 1.  
 XX  
 SQ Sequence 1176 BP; 239 A; 381 C; 354 G; 202 T; 0 other;

Query Match 79.0%; Score 945.2; DB 25; Length 1176;  
 Best Local Similarity 88.3%; Pred. No. 9.8e-191;  
 Matches 1060; Conservative 0; Mismatches 113; Indels 27; Gaps 2;

Qy	1	ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	60
Db	1	ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG	60
Qy	61	CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	120
Db	61	CTGACGCCTGGCCGGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	120
Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTCCGCGCCAGATTCTGTCCAAGCTTCGGCTTGCC	180
Db	121	GTGAAGCGGAAGCGCATCGAGGCCATCCGCGCCAGATCCTGTCCAAGCTGCGGCTCGCC	180
Qy	181	AGCCCCCAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	181	AGCCCCCAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG	240
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG	300
Db	241	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG	300
Qy	301	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACCAAATC	360
Db	301	GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACGAAATC	360



Qy 361 TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420  
 ||||| || ||||| | | ||||| ||||| | ||||| || |||||  
 Db 361 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 420  
 Qy 421 CGGGAAGCGGTGCCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT--GAGG 477  
 || ||||| || ||||| | ||||| ||||| ||||| ||||| |||||  
 Db 421 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGAGG 480  
 Qy 478 CTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGG 537  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 481 CTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGG 540  
 Qy 538 CGCTACCTCAGCAACCGGTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCTTTTGAT 597  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 541 CGATACCTCAGCAACCGGTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGAT 600  
 Qy 598 GTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTC 657  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 601 GTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTT 660  
 Qy 658 AGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTC 717  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 661 AGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTC 720  
 Qy 718 AATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTC 777  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 721 ACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTT 780  
 Qy 778 CTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC 837  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 781 CTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCC 840  
 Qy 838 CTGGATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAG 897  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 841 CTGGACACCAACTATTGCTTCAGCTCCAC-----GGAGAAG 876  
 Qy 898 AACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATT 957  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 877 AACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATC 936  
 Qy 958 CATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGC 1017  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 937 CACGAGCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGC 996  
 Qy 1018 CTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCG 1077  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 997 CTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCG 1056  
 Qy 1078 GCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGC 1137  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1057 GCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGC 1116  
 Qy 1138 CGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1117 CGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1176

Search completed: October 27, 2003, 19:11:02  
Job time : 367.617 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 18:35:27 ; Search time 355.784 Seconds  
(without alignments)  
9022.658 Million cell updates/sec

Title: US-10-017-372E-12  
Perfect score: 1197  
Sequence: 1 atggcgccttcggggctgcg.....gttcctgcaagtgcagctga 1197

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

%

Result	No.	Score	Query Match	Length	DB	ID	Description
	1	958.2	80.1	1821	14	US-10-087-268-1	Sequence 1, Appli
	2	958.2	80.1	2742	14	US-10-037-270-220	Sequence 220, App
	3	956.6	79.9	1821	14	US-10-087-268-4	Sequence 4, Appli
	4	943.6	78.8	2745	11	US-09-948-002-28	Sequence 28, Appl
	5	854.2	71.4	1585	11	US-09-948-002-27	Sequence 27, Appl
	6	847.8	70.8	2094	11	US-09-948-002-1	Sequence 1, Appli
	7	654	54.6	1376	10	US-09-756-283A-19	Sequence 19, Appl
	8	591.4	49.4	1352	10	US-09-756-283A-21	Sequence 21, Appl
	9	342.4	28.6	489	11	US-09-911-904-167	Sequence 167, App
	10	281.2	23.5	339	10	US-09-813-271B-1	Sequence 1, Appli
	11	199.8	16.7	336	10	US-09-813-271B-7	Sequence 7, Appli
	12	198	16.5	2574	11	US-09-906-158-3	Sequence 3, Appli
	13	198	16.5	2574	13	US-10-028-158-20	Sequence 20, Appl
	14	196	16.4	4382	12	US-09-957-458B-9	Sequence 9, Appli
	15	193	16.1	339	10	US-09-813-271B-5	Sequence 5, Appli
	16	190.8	15.9	2879	11	US-09-906-158-10	Sequence 10, Appl
c	17	185.2	15.5	2381	12	US-10-311-455-2384	Sequence 2384, Ap
	18	181.2	15.1	2381	12	US-10-311-455-2383	Sequence 2383, Ap
	19	171.2	14.3	336	10	US-09-813-271B-9	Sequence 9, Appli
	20	166	13.9	336	10	US-09-813-271B-11	Sequence 11, Appl
	21	155	12.9	4267	11	US-09-948-002-47	Sequence 47, Appl
	22	148.6	12.4	339	10	US-09-813-271B-3	Sequence 3, Appli
	23	148.6	12.4	2570	12	US-09-960-706-663	Sequence 663, App
	24	148.6	12.4	2912	13	US-10-044-090-323	Sequence 323, App
c	25	132.6	11.1	597	9	US-09-864-761-15319	Sequence 15319, A
c	26	131.8	11.0	206	10	US-09-833-381-577	Sequence 577, App
c	27	131.8	11.0	537	9	US-09-864-761-8844	Sequence 8844, Ap
c	28	130.4	10.9	148	9	US-09-864-761-25510	Sequence 25510, A
c	29	130.2	10.9	154	9	US-09-864-761-31841	Sequence 31841, A
	30	113.8	9.5	181	10	US-09-833-381-1416	Sequence 1416, Ap
	31	104.8	8.8	851	13	US-10-027-632-152938	Sequence 152938,
c	32	100.4	8.4	224	10	US-09-833-381-73	Sequence 73, Appl
	33	100	8.4	29000	11	US-09-906-158-17	Sequence 17, Appl
	34	73.2	6.1	450	11	US-09-918-995-6145	Sequence 6145, Ap
c	35	69.8	5.8	431	9	US-09-864-761-18644	Sequence 18644, A
	36	69	5.8	584	12	US-10-029-386-9758	Sequence 9758, Ap
	37	68.2	5.7	176	12	US-10-029-386-23458	Sequence 23458, A
c	38	67.8	5.7	176	10	US-09-833-381-571	Sequence 571, App
c	39	64	5.3	363	10	US-09-833-381-585	Sequence 585, App
	40	62.6	5.2	364	12	US-10-029-386-26476	Sequence 26476, A
	41	62.6	5.2	544	12	US-10-029-386-12776	Sequence 12776, A
	42	62.6	5.2	927	12	US-10-244-718-1	Sequence 1, Appli
	43	57.4	4.8	486	11	US-09-918-995-25641	Sequence 25641, A
	44	55.2	4.6	594	13	US-10-027-632-141376	Sequence 141376,
c	45	48.2	4.0	5607	13	US-10-108-605-92	Sequence 92, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-087-268-1

; Sequence 1, Application US/10087268

; Publication No. US20030119010A1

```

; GENERAL INFORMATION:
; APPLICANT: Jonsonn, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or
condition
; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(511)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (512)..(1684)
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (512)..(598)
; OTHER INFORMATION:
; NAME/KEY: 3'UTR
; LOCATION: (1685)..(1821)
; OTHER INFORMATION:
US-10-087-268-1

```

```

Query Match          80.1%;  Score 958.2;  DB 14;  Length 1821;
Best Local Similarity 88.6%;  Pred. No. 1.5e-257;
Matches 1060;  Conservative 0;  Mismatches 113;  Indels 24;  Gaps 1;

```

```

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| |||| || ||||||||||| ||||||||||| ||||||||||| || |||
Db      512 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 571

Qy      61 CTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||
Db      572 CTGACGCCTGGCCGGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 631

Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||
Db      632 GTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 691

Qy      181 AGCCCCCAGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||||||||||| ||||||||||| ||||||||||| |||||| || || || ||
Db      692 AGCCCCCAGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 751

Qy      241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG 300
      |||||| || ||||||||||| |||||| |||| ||||||||||| |||
Db      752 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 811

Qy      301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      || ||||||||||| ||||||||||| ||||||||||| || ||| |||||
Db      812 GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC 871

```

Qy 361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTC 420  
 ||||| || ||||| | | ||||| ||||| | ||||| |||||  
 Db 872 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 931

Qy 421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480  
 || ||||| || ||||| | ||||| ||||| ||||| ||||| |||||  
 Db 932 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 991

Qy 481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC 540  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 992 AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA 1051

Qy 541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTGATGTC 600  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1052 TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTATGTC 1111

Qy 601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCTCAGT 660  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1112 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCTTAGC 1171

Qy 661 GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1172 GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT 1231

Qy 721 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780  
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1232 ACCGGCCGCCGAGGTGACCTGGCCACCATTACATGGCATGAACCGGCCCTTCCTGCTTCTC 1291

Qy 781 ATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 840  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1292 ATGGCCACCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG 1351

Qy 841 GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC 900  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1352 GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC 1387

Qy 901 TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT 960  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1388 TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC 1447

Qy 961 GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA 1020  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1448 GAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCCTG 1507

Qy 1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCCGGCG 1080  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1508 GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCCGGCG 1567

Qy 1081 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1568 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC 1627

Qy 1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAAGTGCAGCTGA 1197  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1628 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTCTTGCAAGTGCAGCTGA 1684

RESULT 2

US-10-037-270-220

```
; Sequence 220, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 220
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (842)..(2014)
```

US-10-037-270-220

```
Query Match          80.1%;  Score 958.2;  DB 14;  Length 2742;
Best Local Similarity 88.6%;  Pred. No. 1.7e-257;
Matches 1060;  Conservative 0;  Mismatches 113;  Indels 24;  Gaps 1;
```

```
Qy      1  ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG  60
      ||| |||| || ||||||||| ||||||||| ||||||||| |||
Db      842 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG  901

Qy      61  CTGACGCCTGGCCGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG  120
      ||||||||| ||||||||| ||||| ||||||||| ||||||||| |||
Db      902 CTGACGCCTGGCCGGCCGGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGGAGCTG  961

Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC  180
      ||||||||| ||||||||| ||||||||| ||||||||| ||||| |||
Db      962 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC  1021
```

Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	1022	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCCGGCCCCGCTGCCCGAGGCCGTGCTCGCCCTG	1081
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCCGAGCCAGAG	300
Db	1082	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCCGAGCCTGAG	1141
Qy	301	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	1142	GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC	1201
Qy	361	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGCGAGCTC	420
Db	1202	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	1261
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	1262	CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	1321
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC	540
Db	1322	AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA	1381
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGATGTC	600
Db	1382	TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTC	1441
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	1442	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	1501
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1502	GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT	1561
Qy	721	TCTGGCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	780
Db	1562	ACCGGCCGCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTC	1621
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1622	ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1681
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1682	GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC	1717
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCAT	960
Db	1718	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1777
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1778	GAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCCTG	1837

```

QY      1021 GACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCGGCG 1080
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1838 GACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCG 1897

QY      1081 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1898 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC 1957

QY      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1958 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 2014

```

RESULT 3

US-10-087-268-4

```

; Sequence 4, Application US/10087268
; Publication No. US20030119010A1
; GENERAL INFORMATION:
; APPLICANT: Jonsonn, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or
condition
; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(511)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (512)..(1684)
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (512)..(598)
; OTHER INFORMATION:
; NAME/KEY: 3'UTR
; LOCATION: (1685)..(1821)
; OTHER INFORMATION:

```

US-10-087-268-4

```

Query Match          79.9%; Score 956.6; DB 14; Length 1821;
Best Local Similarity 88.5%; Pred. No. 4.3e-257;
Matches 1059; Conservative 0; Mismatches 114; Indels 24; Gaps 1;

```

```

QY      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
          ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      512 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 571

QY      61 CTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      572 CTGACGCCTGGCCCGCCGGCCGGCCGACTATCCACCTGCAAGACTATCGACATGGAGCTG 631

```



Qy	121	GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAAGCTTCGGCTTGCC	180
Db	632	GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	691
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGTGCCTGAGGCAGTACTGGCTCTT	240
Db	692	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGTGCCTGAGGCCGTGCTCGCCCTG	751
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCCGAGCCAGAG	300
Db	752	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAAACCGGAGCCCCGAGCCTGAG	811
Qy	301	GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	812	GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC	871
Qy	361	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGCGAGCTC	420
Db	872	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	931
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	932	CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	991
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	540
Db	992	AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	1051
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	600
Db	1052	TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTGATGTC	1111
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	1112	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	1171
Qy	661	GCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1172	GCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACT	1231
Qy	721	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTC	780
Db	1232	ACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGCTTCTC	1291
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1292	ATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG	1351
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1352	GACACCAACTATTGCTTCAGCTCCAC-----GGAGAAGAAC	1387
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	960
Db	1388	TGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCAC	1444

Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1448	GAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCTACATTTGGAGCCTG	1507
Qy	1021	GACACTCAGTACAGCAAGGTCTCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCTGGCG	1080
Db	1508	GACACGCAGTACAGCAAGGTCTCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCTGGCG	1567
Qy	1081	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC	1140
Db	1568	GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGC	1627
Qy	1141	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1197
Db	1628	AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1684

## RESULT 4

US-09-948-002-28

; Sequence 28, Application US/09948002

; Publication No. US20030050265A1

; GENERAL INFORMATION:

; APPLICANT: Nicholas M. Dean

; APPLICANT: Susan F. Murray

TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH

; TITLE OF INVENTION: FACTOR BETA EXPRESSION

; FILE REFERENCE: ISPH-0607

; CURRENT APPLICATION NUMBER: US/09/948,002

; CURRENT FILING DATE: 2000-09-05

; PRIOR APPLICATION NUMBER: 09/661,753

; PRIOR FILING DATE: 2000-09-14

PRIOR APPLICATION NUMBER: 60/154,546

; PRIOR FILING DATE: 1999-09-17

; NUMBER OF SEQ ID NOS: 71

; SEQ ID NO 28

; LENGTH: 2745

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

```
; NAME/KEY: CDS
```

; LOCATION: (842) ... (2017)

US-09-948-002-28

Query Match 78.8%; Score 943.6; DB 11; Length 2745;  
Best Local Similarity 88.2%; Pred. No. 2.le-253;  
Matches 1059; Conservative 0; Mismatches 114; Indels 27; Gaps 2;

[illegible]

Db	962	GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGGCTCGCC	1021
Qy	181	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGGCTCTT	240
Db	1022	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCGCCCTG	1081
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCAGCCAGAG	300
Db	1082	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCAGCCTGAG	1141
Qy	301	GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	1142	GCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC	1201
Qy	361	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGCGAGCTC	420
Db	1202	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	1261
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT---GAGG	477
Db	1262	CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGAGGAGG	1321
Qy	478	CTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTTG	537
Db	1322	CTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGG	1381
Qy	538	CGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCCTTTGAT	597
Db	1382	CGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGAT	1441
Qy	598	GTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTC	657
Db	1442	GTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTT	1501
Qy	658	AGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTC	717
Db	1502	AGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTC	1561
Qy	718	AATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTC	777
Db	1562	ACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCCTGCTT	1621
Qy	778	CTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC	837
Db	1622	CTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCC	1681
Qy	838	CTGGATACCAACTACTGCTTCAGCTCCACGACTACAAGGATGACGACGACAAGGAGAAG	897
Db	1682	CTGGACACCAACTATTGCTTCAGCTCCAC-----GGAGAAG	1717
Qy	898	AACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATT	957
Db	1718	AACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATC	1777
Qy	958	CATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGC	1017
Db	1778	CACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGCCCCCTACATTTGGAGC	1837

```

Qy      1018 CTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCG 1077
          || ||||| ||||||||||||||||||| ||||||||||||||| |||||
Db      1838 CTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCG 1897

Qy      1078 GCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGC 1137
          ||||||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1898 GCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCCGCTGCCCATCGTGTACTACGTGGGC 1957

Qy      1138 CGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
          ||||||||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1958 CGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 2017

```

RESULT 5

US-09-948-002-27

```

; Sequence 27, Application US/09948002
; Publication No. US20030050265A1
; GENERAL INFORMATION:
; APPLICANT: Nicholas M. Dean
; APPLICANT: Susan F. Murray
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH
; TITLE OF INVENTION: FACTOR BETA EXPRESSION
; FILE REFERENCE: ISPH-0607
; CURRENT APPLICATION NUMBER: US/09/948,002
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 09/661,753
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/154,546
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 71
; SEQ ID NO 27
; LENGTH: 1585
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (413)...(1585)
US-09-948-002-27

```

```

Query Match          71.4%; Score 854.2; DB 11; Length 1585;
Best Local Similarity 83.1%; Pred. No. 1.7e-228;
Matches 995; Conservative 0; Mismatches 178; Indels 24; Gaps 1;

```

```

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
          ||| ||||| ||||||||||||||| ||||| ||||| ||| ||| ||||| |||||
Db      413 ATGCCGCCCTCGGGGCTGCGGCTGCTGCCGCTTCTGCTCCCACTCCCGTGGCTTCTAGTG 472

Qy      61 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
          |||||||| || ||||| ||||| ||||| |||||||||||||||||||
Db      473 CTGACGCCCGGGAGGCCAGCCGCGGGACTCTCCACCTGCAAGACCATCGACATGGAGCTG 532

Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
          ||||| ||||||||||||||| ||||| ||||| ||||| ||||| ||| |||
Db      533 GTGAAACGGAAGCGCATCGAAGCCATCCGTGGCCAGATCCTGTCCAACTAAGGCTCGCC 592

Qy      181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240

```

Db	593	ACTCCCCGAGCCAGGGGAGGTACCGCCGGGCCCCGCTGCCCCGAGGCGGTGCTCGCTTTG	652
Qy	241	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCTGAACCGGAGCCCGAGCCAGAG	300
Db	653	TACAACAGCACCCGCGACCGGGTGGCAGGCGAGAGCGCTGACCCGGAGCCCGAGCCCGAG	712
Qy	301	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	713	GCGGACTACTACGCCAAGAAGTCAACCGCGTGCTAATGGTGGACCGCAACAACGCAATC	772
Qy	361	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	420
Db	773	TATGACAAAACCAAAGACATCACACACAGTATATATATGTTCTTCAATACGTCAGACATT	832
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	833	CGGGAAGCAGTGCCAGAACCCCCATTGCTGTCCCGTGACAGAGCTGCGCCTGCAGAGATTC	892
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC	540
Db	893	AAGTCAACTGTGGAGCAACACGTAGAACTCTACCAGAAATATAGCAACAATTCCTGGCGT	952
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGAAGTCAACCGGAGTGGCTGTCTTTGATGTC	600
Db	953	TACCTTGGTAACCGGCTGCTGACCCCCACTGATACGCCTGAGTGGCTGTCTTTTGACGTC	1012
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	1013	ACTGGAGTTGTCCGGCAGTGGCTGAACCAAGGAGACGGAATACAGGGCTTTTCGCTTCAGT	1072
Qy	661	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1073	GCTCACTGCTCTTGTGACAGCAAAGATAATGTACTCCACGTGGAAATCAATGGGATCAGT	1132
Qy	721	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	780
Db	1133	CCCAAACGTCGAGGTGACCTGGGCACCATCCATGACATGAACCGACCTTCCTGCTCCTC	1192
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1193	ATGGCCACCCCCCTGGAAGGGCTCAACACCTGCACAGCTCCAGGCACCGGAGAGCCCTG	1252
Qy	841	GATACCAACTACTGCTTCAGCTCCACGGAATAACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1253	GATACCAACTACTGCTTCAGCTCCAC-----AGAGAAGAAC	1288
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	960
Db	1289	TGCTGTGTACGGCAGCTGTACATTGACTTTAGGAAGGACCTGGGTTGGAAGTGGATCCAC	1348
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1349	GAGCCCAAGGGCTACCATGCCAATTTCTGTCTGGGGCCCTGCCCCTACATTTGGAGCCTG	1408
Qy	1021	GAACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCGGCG	1080

```

Db      1409 GACACACAGTACAGCAAGGTCCTTGCCCTCTACAACCAACACAACCCGGGTGCTTCCGCA 1468
Qy      1081 GCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGC 1140
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1469 TCACCGTGCTGCGTGCCGCAGGCTTTGGAGCCACTGCCCATCGTCTACTACGTGGGTGCG 1528
Qy      1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1529 AAGCCCAAGGTGGAGCAGTTGTCCAACATGATCGTGCGCTCCTGCAAGTGCAGCTGA 1585

```

RESULT 6

US-09-948-002-1

```

; Sequence 1, Application US/09948002
; Publication No. US20030050265A1
; GENERAL INFORMATION:
; APPLICANT: Nicholas M. Dean
; APPLICANT: Susan F. Murray
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH
; TITLE OF INVENTION: FACTOR BETA EXPRESSION
; FILE REFERENCE: ISPH-0607
; CURRENT APPLICATION NUMBER: US/09/948,002
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 09/661,753
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/154,546
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 71
; SEQ ID NO 1
; LENGTH: 2094
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (868)...(2040)

```

US-09-948-002-1

```

Query Match          70.8%; Score 847.8; DB 11; Length 2094;
Best Local Similarity 82.8%; Pred. No. 1.1e-226;
Matches 991; Conservative 0; Mismatches 182; Indels 24; Gaps 1;

```

```

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
          ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      868 ATGCCGCCCTCGGGGCTGCGGCTACTGCCGCTTCTGCTCCCACTCCCGTGGCTTCTAGTG 927
Qy      61 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      928 CTGACGCCCCGGGAGGCCAGCCCGGGACTCTCCACCTGCAAGACCATCGACATGGAGCTG 987
Qy      121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      988 GTGAAACGGAAGCGCATCGAAGCCATCCGTGGCCAGATCCTGTCCAACTAAGGCTCGCC 1047
Qy      181 AGCCCCCCCAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1048 AGTCCCCCAAGCCAGGGGGAGGTACCGCCCCGGCCCCGCTGCCCGAGGCGGTGCTCGCTTTG 1107

```

Qy	241	TACAACAGTACCCCGGACCCGGGTAGCCGGGGAAAAGTGTGGAACCCGAGCCCGAGCCAGAG	300
Db	1108	TACAACAGCACCCCGGACCCGGGTGGCAGGCGAGAGCGCCGACCCAGAGCCCGAGCCCGAA	1167
Qy	301	GCGGACTACTACGCCAAGGAGGTACCCCGCTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	1168	GCGGACTACTATGCTAAAGAGGTACCCCGCTGCTAATGGTGGACCGCAACAACGCCATC	1227
Qy	361	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTC	420
Db	1228	TATGAGAAAACCAAAGACATCTCACACAGTATATATATGTTCTTCAATACGTCCAGACATT	1287
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	1288	CGGGAAGCAGTGCCCGAACCCCATTTGCTGTCCCGTGAGAGCTGCGCTTGAGAGATTA	1347
Qy	481	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAAATACAGCAATGATTCTGGCGC	540
Db	1348	AAATCAAGTGTGGAGCAACATGTGGAACCTTACCAGAAAATAGCAACAATTCTGGCGT	1407
Qy	541	TACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	600
Db	1408	TACCTTGGTAACCGGCTGCTGACCCCCACTGATACGCCCTGAGTGGCTGTCTTTGACGTC	1467
Qy	601	ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	660
Db	1468	ACTGGAGTTGTACGGCAGTGGCTGAACCAAGGAGACGGAATACAGGGCTTTTCGATTGAGC	1527
Qy	661	GCCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	720
Db	1528	GCTCACTGCTCTTGTGACAGCAAAGATAACAACTCCACGTGGAAATCAACGGGATCAGC	1587
Qy	721	TCTGGCCGCCGGGGTGACCTGGCCACCATTACCGCATGAACCGGCCCTTCTGCTCCTC	780
Db	1588	CCCAAACGTCGGGGCGACCTGGGCACCATCCATGACATGAACCGGCCCTTCTGCTCCTC	1647
Qy	781	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	840
Db	1648	ATGGCCACCCCCCTGGAAGGGCCAGCACCTGCACAGCTCACGGCACCGGAGAGCCCTG	1707
Qy	841	GATACCAACTACTGCTTCAGCTCCACGACTACAAGGATGACGACGACAAGGAGAAGAAC	900
Db	1708	GATACCAACTATTGCTTCAGCTCCAC-----AGAGAAGAAC	1743
Qy	901	TGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCAT	960
Db	1744	TGCTGTGTGCGGCAGCTGTACATTGACTTTAGGAAGGACCTGGGTTGGAAGTGGATCCAC	1803
Qy	961	GAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTA	1020
Db	1804	GAGCCCAAGGGCTACCATGCCAATTTCTGTCTGGGACCTGCCCCCTATATTGGAGCCTG	1863
Qy	1021	GACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCAACCCGGGCGCGTCGGCG	1080
Db	1864	GACACACAGTACAGCAAGGTCTTGCCCTCTACAACCAACACAACCCGGGCGCTTCGGCG	1923
Qy	1081	GCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGCG	1140

Db 1924 TCACCGTGCTGCGTGCCG CAGGCTTTGGAGCCACTGCCATCGTCTACTACGTGGGTCGC 1983  
 Qy 1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTCTGCAAGTGCAGCTGA 1197  
 Db 1984 AAGCCCAAGGTGGAGCAGTTGTCCAACATGATTGTGCGTCTCTGCAAGTGCAGCTGA 2040

Qy 1141 AAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 1197

Db 1984 AAGCCCAAGGTGGAGCAGTTGTCCAAATGATTGTGCGCTCCTGCAAGTGCAGCTGA 2040

RESULT 7

US-09-756-283A-19

; Sequence 19, Application US/09756283A

; Patent No. US20020151478A1

; GENERAL INFORMATION:

APPLICANT: Chernajovsky, Yuti

; APPLICANT: Dreja, Hanna Stina

APPLICANT: Adams, Gillian

TITLE OF INVENTION: Latent Fusion Protein

FILE REFERENCE: 0623.1000000

CURRENT APPLICATION NUMBER: US/09/756,283A

; CURRENT FILING DATE: 2001-01-09

NUMBER OF SEQ ID NOS: 100

```

; SOFTWARE: PatentIn version 3.0

```

; SEQ ID NO 19

: LENGTH: 1376

; TYPE: DNA

; ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: LAP-mIFNbeta construct

; NAME/KEY: CDS

```

; LOCATION: (1)..(1368)

```

US-09-756-283A-19

Query Match 54.6%; Score 654; DB 10; Length 1376;

Best Local Similarity 88.1%; Pred. No. 1.4e-172;

Matches 724; Conservative 0; Mismatches 95; Indels 3; Gaps 1;

Qv 1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60

Db 1 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 60

Qy 61 CTGACGCCTGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120

Db 61 CTGACGCCTGGCCCGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 120

Qy 121 GTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180

Db 121 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 180

QY 181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTT 240

Db 181 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 240

Qy 241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGCCAGAG 300

Db 241 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 300

Qv 301 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360



```

      || |||||
Db      301 GCCGACTACTACGCCAAGGAGGTCACCCGCTGCTAATGGTGGAAACCCACAACGAAATC 360
Qy      361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420
      ||||| || ||||| || ||||| ||||| ||||| ||||| |||||
Db      361 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 420
Qy      421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT--GAGG 477
      || ||||| || ||||| || ||||| ||||| ||||| ||||| |||||
Db      421 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGAGG 480
Qy      478 CTCAAGTTAAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCTGG 537
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      481 CTCAAGTTAAAAAGTGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGG 540
Qy      538 CGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGAT 597
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      541 CGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGAT 600
Qy      598 GTCACCGGAGTTGTGCGGCAGTGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCCTC 657
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      601 GTCACCGGAGTTGTGCGGCAGTGCTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTT 660
Qy      658 AGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTC 717
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      661 AGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTC 720
Qy      718 AATTCTGGCCCGCCGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCCTGCTC 777
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      721 ACTACCGGCCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCCTGCTT 780
Qy      778 CTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGC 819
      ||||| ||||| ||||| ||||| ||||| |||||
Db      781 CTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGC 822

```

RESULT 8

```

US-09-756-283A-21
; Sequence 21, Application US/09756283A
; Patent No. US20020151478A1
; GENERAL INFORMATION:
; APPLICANT: Chernajovsky, Yuti
; APPLICANT: Dreja, Hanna Stina
; APPLICANT: Adams, Gillian
; TITLE OF INVENTION: Latent Fusion Protein
; FILE REFERENCE: 0623.1000000
; CURRENT APPLICATION NUMBER: US/09/756,283A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 1352
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mIFNbeta-LAP construct
; NAME/KEY: CDS

```

US-09-756-283A-21

Query Match 49.4%; Score 591.4; DB 10; Length 1352;  
Best Local Similarity 87.0%; Pred. No. 4.1e-155;  
Matches 662; Conservative 0; Mismatches 96; Indels 3; Gaps 1;

Qy 63 GACGCTGGCCGCCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTGGT 122  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 582 GGCGGGAGGGGGCTCAGCGGCCGCACCTATCCACCTGCAAGACTATCGACATGGAGCTGGT 641

Qy 123 GAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAG 182  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 642 GAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAG 701

Qy 183 CCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTA CTGGCTCTTTA 242  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 702 CCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTGTA 761

Qy 243 CAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCAGCCAGAGGC 302  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 762 CAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCAGCCTGAGGC 821

Qy 303 GGA TACTACGCCAAGGAGGT CACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTA 362  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 822 CGACTACTACGCCAAGGAGGT CACCCGCGTGCTAATGGTGGAAACCCACAACGAAATCTA 881

Qy 363 TGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTCCG 422  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 882 TGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCG 941

Qy 423 GGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT - - -GAGGCT 479  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 942 AGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGAGGCT 1001

Qy 480 CAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCG 539  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1002 CAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCG 1061

Qy 540 CTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCCTTTGATGT 599  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1062 ATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTG GTTATCTTTTGATGT 1121

Qy 600 CACCGGAGTTGTGCGGCAGTGGCTGACCCG CAGAGAGGCTATAGAGGGTTTTCGCCTCAG 659  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1122 CACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAG 1181

Qy 660 TGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAA 719  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1182 CGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCA 1241

Qy 720 TTCTGGCCGCCCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCCTGCTCCT 779  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1242 TACCGGCCGCCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCTTCCTGCTTCT 1301

Qy 780 CATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCT 820  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 1302 CATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCT 1342

Query Match 28.6%; Score 342.4; DB 11; Length 489;  
Best Local Similarity 89.6%; Pred. No. 1.1e-85;  
Matches 389; Conservative 0; Mismatches 21; Indels 24; Gaps 1;

Qy 1124 TGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC GTTCCT 1183  
 |||  
 Db 337 TGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCTGAACATGATCGTGC GTCTCCT 396  
 Qy 1184 GCAAGTGCAGCTGA 1197  
 |||  
 Db 397 GCAAGTGCAGCTGA 410

RESULT 10

US-09-813-271B-1

; Sequence 1, Application US/09813271B

; Patent No. US20020115834A1

; GENERAL INFORMATION:

; APPLICANT:

; (A) Nico Cerletti

; TITLE OF INVENTION: New process for the production of  
 ; biologically active protein

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. US20020115834A1artis Patent Department

; STREET: 564 Morris Avenue

; CITY: Summit

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07901

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/813,271B

; FILING DATE: 20-Mar-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP95/02719

; FILING DATE: 12-Jul-95

; APPLICATION NUMBER: EPO 94810439.3

; FILING DATE: 25-Jul-94

; ATTORNEY/AGENT INFORMATION:

; NAME: Pfeiffer, Hesna J. .

; REGISTRATION NUMBER: 22640

; REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (908) 522-6940

; TELEFAX: (908) 522-6955

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 339 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; HYPOTHETICAL: NO

; IMMEDIATE SOURCE:

; CLONE: E. coli LC137/pPLMu.hTGF-beta1 (DSM 5656)

```

;         FEATURE:
;         NAME/KEY:  CDS
;         LOCATION:  1..336
;         OTHER INFORMATION:  /product= "human TGF-beta1"
;         SEQUENCE DESCRIPTION:  SEQ ID NO: 1:
US-09-813-271B-1

```

```

Query Match          23.5%;  Score 281.2;  DB 10;  Length 339;
Best Local Similarity  94.2%;  Pred. No. 1.3e-68;
Matches 292;  Conservative  0;  Mismatches 18;  Indels  0;  Gaps  0;

```

```

Qy      888 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 947
      || |||||
Db      30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Qy      948 GAAGTGGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 1007
      |||||
Db      90 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 149

Qy     1008 CATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCC 1067
      |||
Db     150 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCC 209

Qy     1068 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 1127
      |||||
Db     210 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCCGCTGCCCATCGTGTA 269

Qy     1128 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTGCAA 1187
      |||||
Db     270 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTGCAA 329

Qy     1188 GTGCAGCTGA 1197
      |||||
Db     330 GTGCAGCTGA 339

```

RESULT 11

US-09-813-271B-7

```

; Sequence 7, Application US/09813271B
; Patent No. US20020115834A1
;   GENERAL INFORMATION:
;     APPLICANT:
;       (A) Nico Cerletti
;     TITLE OF INVENTION: New process for the production of
;                           biologically active protein
;     NUMBER OF SEQUENCES: 13
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: No. US20020115834A1artis Patent Department
;       STREET: 564 Morris Avenue
;       CITY: Summit
;       STATE: New Jersey
;       COUNTRY: USA
;       ZIP: 07901
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

;      SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/813,271B
;      FILING DATE: 20-Mar-2001
;
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: PCT/EP95/02719
;      FILING DATE: 12-Jul-95
;      APPLICATION NUMBER: EPO 94810439.3
;      FILING DATE: 25-Jul-94
;
;      ATTORNEY/AGENT INFORMATION:
;      NAME: Pfeiffer, Hesna J. .
;      REGISTRATION NUMBER: 22640
;      REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN
;
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (908) 522-6940
;      TELEFAX: (908) 522-6955
;
;      INFORMATION FOR SEQ ID NO: 7:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 336 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;
;      MOLECULE TYPE: other nucleic acid
;      DESCRIPTION: /desc = "recombinant hybrid DNA of
;      IMMEDIATE SOURCE:
;      CLONE: E. coli LC137/pPLMu.TGF-beta1(44/45)beta3
;
;      FEATURE:
;      NAME/KEY: mat_peptide
;      LOCATION: 1..132
;      OTHER INFORMATION: /product= "N-terminal 44 amino
;      acids of human TGF-beta1"
;
;      FEATURE:
;      NAME/KEY: mat_peptide
;      LOCATION: 133..336
;      OTHER INFORMATION: /product= "C-terminal 68 amino
;      acids of human TGF-beta3"
;
;      FEATURE:
;      NAME/KEY: CDS
;      LOCATION: 1..336
;      OTHER INFORMATION: /product= "hybrid TGF-beta named
;      TGF-beta1-3"
;
;      SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-813-271B-7

```

```

Query Match          16.7%;  Score 199.8;  DB 10;  Length 336;
Best Local Similarity 78.2%;  Pred. No. 7e-46;
Matches 240;  Conservative 0;  Mismatches 67;  Indels 0;  Gaps 0;

```

```

Qy      888 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 947
      || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30  CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Qy      948 GAAGTGGATTTCATGAACCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 1007
      ||||||||| || || |||||||||||||||||||||||||||| || || || || ||
Db      90  GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCTCAGGCCCTTGCCCAT 149

Qy      1008 CATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGGCTCTGTACAACCAGCACAAACC 1067

```

```

      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      150 CCTCCGAGTGCAGACACAACCCACAGCACGGTGCTGGGACTGTACAACACTCTGAACCC 209
Qy      1068 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 1127
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      210 TGAAGCATCTGCCTCGCCTTGCTGCGTGCCCCAGGACCTGGAGCCCCTGACCATCCTGTA 269
Qy      1128 CTACGTGGGCGCAAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 1187
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      270 CTATGTTGGGAGGACCCCCAAAGTGGAGCAGCTCTCCAACATGGTGGTGAAGTCTTGTA 329
Qy      1188 GTGCAGC 1194
      | | | |
Db      330 ATGTAGC 336

```

RESULT 12

US-09-906-158-3

; Sequence 3, Application US/09906158

; Publication No. US20030078217A1

; GENERAL INFORMATION:

; APPLICANT: Brett P. Monia

; APPLICANT: Susan M. Freier

; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH FACTOR-BETA 3  
EXPRESSION

; FILE REFERENCE: RTS-0257

; CURRENT APPLICATION NUMBER: US/09/906,158

; CURRENT FILING DATE: 2001-07-14

; NUMBER OF SEQ ID NOS: 168

; SEQ ID NO 3

; LENGTH: 2574

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (254)...(1492)

US-09-906-158-3

Query Match 16.5%; Score 198; DB 11; Length 2574;

Best Local Similarity 51.0%; Pred. No. 3.9e-45;

Matches 631; Conservative 0; Mismatches 535; Indels 72; Gaps 4;

```

Qy      26 TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCCG 85
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      261 TGCACTTGCAAAGGGCTCTGGTGGTCCTGGCCCTGCTGAACTTTGCCACGGTCAGCCTCT 320
Qy      86 GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 145
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      321 CTCTGTCCACTTGCAACACCTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGAAGCCA 380
Qy      146 TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 205
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      381 TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCAACGGTGATGA 440
Qy      206 CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 265
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      441 CCCACG-----TCCCCTATCAGGTCCTGGCCCTTTACAACAGCACCCGGGAGCTGCTGG 494

```

Qy	266	CCGGGGAAAGTGTGCGAACCGGAGCCCG-----AGCCAGAGGCGGACTACT	310
Db	495	AGGAGATGCATGGGGAGAGGGAGGAAGGCTGCACCCAGGAAAACACCGAGTCGGAATACT	554
Qy	311	ACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAT	370
Db	555	ATGCCAAAGAAATCCATAAATTCGACATGATCCAGGGGCTGGCGGAGCACAAACGAACCTGG	614
Qy	371	TCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGG	430
Db	615	CTGTCTGCCCTAAAGGAATTACCTCCAAGGTTTTCCGCTTCAATGTGTCCTCAGTGGAGA	674
Qy	431	TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAG	490
Db	675	AAAATAGAACCAACCTATTCCGAGCAGAATTCGGGTCTTGCGGTGCCCAACCCAGCT	734
Qy	491	TGGAGCAGCACGTGGAG-----CTATACCAGAAATACAGCAATG	529
Db	735	CTAAGCGGAATGAGCAGAGGATCGAGCTCTTCCAGATCCTTCGGCCAGATGAGCACATTG	794
Qy	530	ATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGT	589
Db	795	CCAAACAGCGCTATATCGGTGGCAAGAATCTGCCACACGGGGCACTGCCGAGTGGCTGT	854
Qy	590	CCTTTGATGTACCCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTT	649
Db	855	CCTTTGATGTCACTGACACTGTGCGTGAGTGGCTGTTGAGAAGAGAGTCCAACCTTAGGTC	914
Qy	650	TTCGCCTCAGTGCCCACTCTTCTCT-----GACA	679
Db	915	TAGAAATCAGCATTCACTGTCCATGTCAACCTTTCAGCCCAATGGAGATATCCTGGAAA	974
Qy	680	GCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCATTCTGGCCGCCGGGTGACC	739
Db	975	ACATTACAGAGGTGATGGAAATCAAATTCAAAGGCGTGGACAATGAGGATGACCATGGCC	1034
Qy	740	TGGCCACCATTCACGGCATGAACCGGCCCTTCTTGCTCCTCATGGCCACCCGCTGGAGA	799
Db	1035	GTGGAGATCTGGGGCGCCTCAAGAAGCAGAAGGATCACCACAACCCTCATCTAATCCTCA	1094
Qy	800	GGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACTACTGCTTCA	859
Db	1095	TGATGATTCCCCACACCGGCTCGACAACCGGGCCAGGGGGGTGAGAGGAAGAAGCGGG	1154
Qy	860	GCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAAGTGTGCGTGCGGCAGCTCT	919
Db	1155	CTTTGGACACCAATTACTGCTTCCGCAACTTGGAGGAGAAGTGTGTGTGCGCCCCCTCT	1214
Qy	920	ACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATG	979
Db	1215	ACATTGACTTCCGACAGGATCTGGGCTGGAAGTGGGTCCATGAACCTAAGGGCTACTATG	1274
Qy	980	CCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGG	1039
Db	1275	CCAATTCTGCTCAGGCCCTTGCCCATACCTCCGCAGTGCAGACACAACCCACAGCACGG	1334



```

Qy      1040 TCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGC 1099
          |  ||||  ||||| ||||  |  ||||  |  ||  ||  ||  ||||  ||||| ||||| ||
Db      1335 TGCTGGGACTGTACAACACTCTGAACCCTGAAGCATCTGCCTCGCCTTGCTGCGTGCCCC 1394

Qy      1100 AGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGCGCGCAAGCCCCAAGGTGGAGCAGC 1159
          |||  ||||| ||||  ||||  ||||| ||||| ||  ||  ||  ||||  ||||| |||||
Db      1395 AGGACCTGGAGCCCCTGACCATCCTGTACTATGTTGGGAGGACCCCCAAAGTGGAGCAGC 1454

Qy      1160 TGTCCAACATGATCGTTCCTGCAAGTGCAGCTGA 1197
          |  ||||| ||||  |||  ||  ||  ||  |||||
Db      1455 TCTCCAACATGGTGGTGAAGTCTTGTAATGTAGCTGA 1492

```

# RESULT 13

US-10-028-158-20

```

; Sequence 20, Application US/10028158
; Publication No. US20020110833A1
; GENERAL INFORMATION:
; APPLICANT: Caniggia, Isabella
; APPLICANT: Post, Martin
; APPLICANT: Lye, Stephen
; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF
; TITLE OF INVENTION: TROPHOBLAST
; FILE REFERENCE: 11757.38USWO
; CURRENT APPLICATION NUMBER: US/10/028,158
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US/09/380,662
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: PCT/CA98/00180
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 60/039,919
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (254)..(1492)
US-10-028-158-20

```

```

Query Match          16.5%;  Score 198;  DB 13;  Length 2574;
Best Local Similarity 51.0%;  Pred. No. 3.9e-45;
Matches 631;  Conservative 0;  Mismatches 535;  Indels 72;  Gaps 4;

```

```

Qy      26  TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCCG 85
          |||  |||  ||  ||| ||  |||  |||||  ||  |  |  |  |
Db      261 TGCACTTGCAAAGGGCTCTGGTGGTCTGGCCCTGCTGAACCTTGCCACGGTCAGCCTCT 320

Qy      86  GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 145
          ||||| ||||  |||  |||  ||  |  |  |  |||  ||||  |  ||  ||||
Db      321 CTCTGTCCACTTGCAACACCTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGAAGCCA 380

Qy      146 TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 205
          ||  ||  |||||  ||  |||||  ||||  ||||| ||||  |  ||  ||

```

Db	381	TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCAACGGTGATGA	440
Qy	206	CGCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG	265
Db	441	CCCACG-----TCCCCTATCAGGTCCTGGCCCTTTACAACAGCACCCGGGAGCTGCTGG	494
Qy	266	CCGGGGAAAGTGTCTGAACCGGAGCCCC-----AGCCAGAGGCGGACTACT	310
Db	495	AGGAGATGCATGGGGAGAGGGAGGAAGGCTGCACCCAGGAAAACACCGAGTCGGAATACT	554
Qy	311	ACGCCAAGGAGGTCAACCGCGTGCTAATGGTGAAAGCGGCAACCAAATCTATGATAAAT	370
Db	555	ATGCCAAAGAAATCCATAAATTCGACATGATCCAGGGGCTGGCGGAGCACAACGAACCTGG	614
Qy	371	TCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGG	430
Db	615	CTGTCTGCCCTAAAGGAATTACCTCCAAGGTTTTCCGCTTCAATGTGTCCTCAGTGGAGA	674
Qy	431	TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAG	490
Db	675	AAAATAGAACCAACCTATTCCGAGCAGAATTCCGGGCTTTGCGGGTGCCCAACCCAGCT	734
Qy	491	TGGAGCAGCACGTGGAG-----CTATACCAGAAATACAGCAATG	529
Db	735	CTAAGCGGAATGAGCAGAGGATCGAGCTCTTCCAGATCCTTCGGCCAGATGAGCACATTG	794
Qy	530	ATTCTTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGT	589
Db	795	CCAAACAGCGCTATATCGGTGGCAAGAATCTGCCCACACGGGGCACTGCCGAGTGGCTGT	854
Qy	590	CCTTTGATGTCAACGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTT	649
Db	855	CCTTTGATGTCACTGACACTGTGCGTGAGTGGCTGTTGAGAAGAGAGTCCAACCTTAGGTC	914
Qy	650	TTCGCCTCAGTGCCCACTCTTCCTCT-----GACA	679
Db	915	TAGAAATCAGCATTCACTGTCCATGTACACCTTTCAGCCCAATGGAGATATCCTGGAAA	974
Qy	680	GCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACC	739
Db	975	ACATTACAGAGGTGATGGAATCAAATTCAAAGGCGTGACAAATGAGGATGACCATGGCC	1034
Qy	740	TGGCCACCATTACGGCATGAACCGGCCCTTCTTGCTCCTCATGGCCACCCGCTGGAGA	799
Db	1035	GTGGAGATCTGGGGCGCCTCAAGAAGCAGAAGGATCACCACAACCTCATCTAATCCTCA	1094
Qy	800	GGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACTACTGCTTCA	859
Db	1095	TGATGATTCCCCCACACCGGCTCGACAACCCGGGCCAGGGGGGTGAGAGGAAGAAGCGGG	1154
Qy	860	GCTCCACGGAATAAAGGATGACGACGACAAGGAGAAAGAACTGCTGCGTGCGGCAGCTCT	919
Db	1155	CTTTGGACACCAATTACTGCTTCGGCAACTTGAGGAGAACTGCTGTGTGCGCCCCCTCT	1214
Qy	920	ACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATG	979
Db	1215	ACATTGACTTCCGACAGGATCTGGGCTGGAAGTGGGTCCATGAACCTAAGGGCTACTATG	1274

Qy	980	CCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGG	1039
Db	1275	CCAACTTCTGCTCAGGCCCTTGCCCATACCTCCGCAGTGCAGACACAACCCACAGCACGG	1334
Qy	1040	TCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCTGGCGGCGCCGTGCTGCGTGCCGC	1099
Db	1335	TGCTGGGACTGTACAACACTCTGAACCTGAAGCATCTGCCTCGCCTTGCTGCGTGCCCC	1394
Qy	1100	AGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGC	1159
Db	1395	AGGACCTGGAGCCCTGACCATCCTGTACTATGTTGGGAGGACCCCAAGTGGAGCAGC	1454
Qy	1160	TGTCCAACATGATCGTGC GTTCCTGCAAGTGCAGCTGA	1197
Db	1455	TCTCCAACATGGTGGTGAAGTCTGTGAAATGTAGCTGA	1492

```

Query Match          16.4%;  Score 196;  DB 12;  Length 4382;
Best Local Similarity 51.8%;  Pred. No. 1.6e-44;
Matches 646;  Conservative 0;  Mismatches 530;  Indels 72;  Gaps 6;

Qy      16  CTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCC'TGGCCGG 75
          ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      455  CTGCAGCCCATGCACTTGCAAAGGGCTCTGGTAGTCCTGGCCCTGCTGAACTTGGCCACA 514

Qy      76  CCGGCCGCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGC 135
          | | ||||| |||| | | ||| | | ||| ||| ||| |
Db      515  ATCAGCCTCTCTCTGTCCACTTGCAACACGTGGACTTCGGCCACATCAAGAAGAAGAGG 574

Qy     136  ATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCCCAGCCAG 195

```

Db	575	GTGGAAGCCATTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCA	634
Qy	196	GGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGC	255
Db	635	TCGGTGATGACCCACG-----TCCCCTATCAGGTCCTGGCACTTTACAACAGCACCCGG	688
Qy	256	GACCGGGTAGCCG-----GGGAAAGTGTCGAACCGGAGCCCCGAGCCAGAG	300
Db	689	GAGTTGCTGGAAGAGATGCACGGGGAGAGGGAGGAAGGCTGCACTCAGGAGACCTCGGAG	748
Qy	301	GCGGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	360
Db	749	TCTGAGTACTATGCCAAAGAGATCCATAAATTTCGACATGATCCAGGGACTGGCGGAGCAC	808
Qy	361	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCGAGCTC	420
Db	809	AATGAACTGGCCGTCTGCCCAAAGGAATTACCTCTAAGGTTTTTCGTTTCAATGTGTCC	868
Qy	421	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	480
Db	869	TCAGTGGAGAAAAATGGAACCAATCTGTTCCGGGCAGAGTTCGGGTCTTGCGGGTGCCC	928
Qy	481	A-----AGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAA	519
Db	929	AACCCAGCTCCAAGCGCACAGAGCAGAGAATTGAGCTCTTCCAGATACTTCGACCGGAT	988
Qy	520	TACAGCAATGATTCTTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCG	579
Db	989	GAGCACATAGCCAAGCAGCGCTACATAGGTGGCAAGAATCTGCCACAAGGGGCACCGCT	1048
Qy	580	GAGTGGCTGTCCTTTGATGTCAACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCT	639
Db	1049	GAATGGCTGTCTTTCGATGTCACTGACACTGTGCGCGAGTGGCTGTTGAGGAGAGAGTCC	1108
Qy	640	ATAGAGGGTTTTTCGCCTCAGTGCCCACTCTTCTCTGACAGC-----AA	683
Db	1109	AACCTTGGGTCTGGAAATCAGCATCCACTGTCCATGTCAACCTTTCAGCCCAATGGAGAC	1168
Qy	684	AGATAACACACTCCACGTGGAAATTAACGGGTTCATTCTGGCCGCCGGGTGACCTGGC	743
Db	1169	ATACTGGAAAAATGTTTCATGAGGTGATGGAAATCAAATTCAAAGGAGTGGACAATGAAGAT	1228
Qy	744	CACCATTCACGGCATGAACCGGCCCTTCT-----GCTCCTCATGGCCACCCCG	792
Db	1229	GACCATGGCCGTGGAGACCTGGGGCGTCTCAAGAAGCAAAAGGATCACCAACCCACAC	1288
Qy	793	CTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACTAC	852
Db	1289	CTGATCCTCATGATGATCCCCCACACCGACTGGACAGCCAGGCCAGGGCAGTCAGAGG	1348
Qy	853	TGCTTCAGCTCCACGGACTACAAGGATGACGACGACAA--GGAGAAGAACTGCTGCGTG	909
Db	1349	AAGAAGAGGGCCCTGGACACCAATTACTGCTTCCGCAACCTGGAGGAGAACTGCTGTGTA	1408
Qy	910	CGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAG	969

Db	1409	CGCCCCCTTTATATTGACTTCCGGCAGGATCTAGGCTGGAAATGGGTCCACGAACCTAAG	1468
Qy	970	GGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAG	1029
Db	1469	GGTTACTATGCCAACTTCTGCTCAGGCCCTTGCCCATACCTCCGAGCGCAGACACAACC	1528
Qy	1030	TACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCTCGGCGGCGCCGTGC	1089
Db	1529	CATAGCACGGTGCTTGGACTATACAACACCCTGAACCCAGAGGCGTCTGCCTCGCCATGC	1588
Qy	1090	TGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTIONTACGTGGGCCGCAAGCCCAAG	1149
Db	1589	TGCGTCCCCCAGGACCTGGAGCCCCTGACCATCTTGTACTIONTATGTGGGCAGAACCCCAAG	1648
Qy	1150	GTGGAGCAGCTGTCCAACATGATCGTTCGTTCCCTGCAAGTGCAGCTGA	1197
Db	1649	GTGGAGCAGCTGTCCAACATGGTGGTGAAGTCGTGTAAGTGCAGCTGA	1696

RESULT 15

; Sequence 5, Application US/09813271B

; Patent No. US20020115834A1

; GENERAL INFORMATION:

APPLICANT:

;

;

(A) Nico Cerletti

TITLE OF INVENTION: New process for the production of  
biologically active protein

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. US20020115834A1artis Patent Department

STREET: 564 Morris Avenue

CITY: Summit

STATE: New Jersey

COUNTRY: USA

ZIP: 07901

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

```

;

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/813,271B

FILING DATE: 20-Mar-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP95/02719

FILING DATE: 12-Jul-95

APPLICATION NUMBER: EPO 94810439.3

FILING DATE: 25-Jul-94

ATTORNEY/AGENT INFORMATION:

NAME: Pfeiffer, Hesna J. .

REGISTRATION NUMBER: 22640

REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN

; TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 522-6940

TELEFAX: (908) 522-6955

INFORMATION FOR SEQ ID NO: 5:

```

; SEQUENCE CHARACTERISTICS:
;   LENGTH: 339 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; IMMEDIATE SOURCE:
;   CLONE: E. coli LC137/pPLMu.hTGF-beta3 (DSM 5658)
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..336
;   OTHER INFORMATION: /product= "human TGF-beta3"
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-813-271B-5

```

```

Query Match          16.1%;  Score 193;  DB 10;  Length 339;
Best Local Similarity 76.0%;  Pred. No. 5.6e-44;
Matches 238;  Conservative 0;  Mismatches 75;  Indels 0;  Gaps 0;

```

```

Qy      885 CGACAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGG 944
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      27  CAACTTGGAGGAGAAGAACTGCTGTGTGCGCCCCCTCTACATTGACTTCCGACAGGATCTGGG 86

Qy      945 CTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCC 1004
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      87  CTGGAAGTGGGTCCATGAACCTAAGGGCTACTATGCCAATTCTGCTCAGGCCCTTGCCC 146

Qy      1005 CTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAA 1064
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      147 ATACCTCCGCAGTGCGAGACACAACCCACAGCACGGTGCTGGGACTGTACAACACTCTGAA 206

Qy      1065 CCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCACTGCCCATCGT 1124
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      207 CCCTGAAGCATCTGCCTCGCCTTGCTGCGTGCCCCAGGACCTGGAGCCCCTGACCATCCT 266

Qy      1125 GTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTG 1184
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      267 GTACTATGTTGGGAGGACCCCCAAAGTGGAGCAGCTCTCCAACATGGTGGTGAAGTCTTG 326

Qy      1185 CAAGTGCAGCTGA 1197
        || || ||| |||
Db      327 TAAATGTAGCTGA 339

```

Search completed: October 28, 2003, 09:04:14  
Job time : 361.927 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 10:47:27 ; Search time 2970.42 Seconds  
(without alignments)  
9794.056 Million cell updates/sec

Title: US-10-017-372E-12  
 Perfect score: 1197  
 Sequence: 1 atggcgcccttcggggctgcg.....gttcctgcaagtgcagctga 1197

Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	%	Query				
No.	Score	Match Length DB ID				Description

	1	717.6	59.9	1072	12	BM562135	BM562135	AGENCOURT
	2	702	58.6	900	13	BX349319	BX349319	BX349319
c	3	692.2	57.8	1201	13	BX355682	BX355682	BX355682
c	4	685.8	57.3	983	13	BX335351	BX335351	BX335351
c	5	629.6	52.6	1041	13	BX324511	BX324511	BX324511
	6	629.4	52.6	1043	12	BM555996	BM555996	AGENCOURT
c	7	603.6	50.4	950	13	BX383773	BX383773	BX383773
c	8	589	49.2	888	13	BX434425	BX434425	BX434425
	9	577.2	48.2	859	9	AL530081	AL530081	AL530081
	10	570.6	47.7	713	10	BE312000	BE312000	601154768
	11	562.6	47.0	717	10	BE260971	BE260971	601153715
	12	561.6	46.9	902	13	BQ675698	BQ675698	AGENCOURT
	13	534.2	44.6	925	12	BI818841	BI818841	603037307
	14	521.2	43.5	1093	12	BQ054305	BQ054305	AGENCOURT
c	15	520.2	43.5	841	9	AL530080	AL530080	AL530080
c	16	520	43.4	956	12	BI084718	BI084718	602869722
	17	514	42.9	1049	12	BQ053417	BQ053417	AGENCOURT
	18	513.6	42.9	871	13	BQ952138	BQ952138	AGENCOURT
	19	508.4	42.5	773	12	BI195242	BI195242	602944518
	20	506.4	42.3	1013	12	BM469326	BM469326	AGENCOURT
	21	504.6	42.2	943	12	BI909079	BI909079	603070060
	22	501.8	41.9	821	10	BG748049	BG748049	602705544
	23	486.8	40.7	637	12	BG938856	BG938856	cn28g07.y
	24	482.8	40.3	748	10	BF682938	BF682938	602117568
c	25	482.2	40.3	757	14	CA309731	CA309731	UI-H-FT1-
	26	475.6	39.7	565	14	CD287835	CD287835	1_M15.abd
c	27	474.8	39.7	1047	13	BX325288	BX325288	BX325288
	28	474.6	39.6	659	10	BG500488	BG500488	602544826
	29	472.8	39.5	675	10	BE261764	BE261764	601147839
	30	464.2	38.8	960	13	BU957001	BU957001	AGENCOURT
	31	460.6	38.5	814	12	BI686959	BI686959	603313837
	32	460.2	38.4	599	12	BI905963	BI905963	603062849
c	33	455.2	38.0	727	9	AI131171	AI131171	qc15c11.x
	34	453.4	37.9	633	14	CB577186	CB577186	AMGNNUC:N
c	35	453.2	37.9	785	9	AI148173	AI148173	qb56d01.x
	36	452.2	37.8	548	12	BM068788	BM068788	ie85a02.y
	37	449.8	37.6	982	13	BQ672677	BQ672677	AGENCOURT
	38	447.4	37.4	981	12	BM802192	BM802192	AGENCOURT
	39	445.8	37.2	588	13	BX097979	BX097979	BX097979
c	40	445.6	37.2	928	13	BX412434	BX412434	BX412434
	41	442.4	37.0	636	10	BE615330	BE615330	601280728
c	42	439.6	36.7	722	13	BU632686	BU632686	UI-H-FE1-
c	43	439.4	36.7	1015	10	BG683840	BG683840	602651694
c	44	437.6	36.6	695	14	CA425775	CA425775	UI-H-FE1-
c	45	436.6	36.5	700	13	BU633922	BU633922	UI-H-FL1-

#### ALIGNMENTS

#### RESULT 1

BM562135

LOCUS BM562135 1072 bp mRNA linear EST 20-FEB-2002  
 DEFINITION AGENCOURT\_6562032 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5745463  
 5', mRNA sequence.  
 ACCESSION BM562135



```

VERSION      BM562135.1  GI:18807966
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 1072)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM12768  row: h  column: 08
              High quality sequence start: 9
              High quality sequence stop: 647.

FEATURES             Location/Qualifiers
     source            1..1072
                       /organism="Homo sapiens"
                       /mol_type="mRNA"
                       /db_xref="taxon:9606"
                       /clone="IMAGE:5745463"
                       /tissue_type="leukocyte"
                       /lab_host="DH10B"
                       /clone_lib="NIH_MGC_118"
                       /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                       (destroyed); RNA source leukocytes from anonymous pool of
                       non-activated adult donors. Library is oligo-dT primed
                       and directionally cloned (EcoRV site is destroyed upon
                       cloning). Average insert size 1.7 kb, insert size range
                       1.2-3.3 kb. Library is normalized and enriched for
                       full-length clones and was constructed by C. Gruber
                       (Invitrogen). Research Genetics tracking code 027. Note:
                       this is a NIH MGC Library."

BASE COUNT      217 a      356 c      313 g      186 t
ORIGIN

Query Match          59.9%;  Score 717.6;  DB 12;  Length 1072;
Best Local Similarity 83.5%;  Pred. No. 1.1e-151;
Matches 885;  Conservative 0;  Mismatches 144;  Indels 31;  Gaps 5;

Qy      117  GCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCT 176
          |||
Db      20   GCTGGTGAAGCGGAAGCGCATCTAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCT 79

Qy      177  TGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGC 236
          |||
Db      80   CGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGC 139

Qy      237  TCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCCGAGCC 296
          |||

```

Db	140 CCTGTACAACAGCACCCGCGACC GG GTGGCCCGGGAGAGTGCAGAACCGGAGCCC GAGCC	199
Qy	297 AGAGGCGGACTACTACGCCAAGGAGGT C ACC CGCTGCTAATGGTGGA AAGC GGCAACCA	356
Db	200 TGAGGCCGACTACTACGCCAAGGAGGT C ACC CGCTGCTAATGGTGGA AACCCA CAACGA	259
Qy	357 AATCTATGATAAAATTCAAGGGCACCCCCACAGCTT ATATATGCTGTTCAACACGTCGGA	416
Db	260 AATCTATGACAAGTTCAAGCAGAGT ACACA CAGCATATATATGTTCTTCAACACATCAGA	319
Qy	417 GCTCCGGGAAGCGGTGCCGGAACCTGT ATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAG	476
Db	320 GCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAG	379
Qy	477 GCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATA CCAGAAATACAGCAATGATTCTCTG	536
Db	380 GCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAA CAATTCTCTG	439
Qy	537 GCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTCTTTGA	596
Db	440 GCGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTTTGA	499
Qy	597 TGTCAACCGGAGTTGTGCGGCAGTGGCTGACCCG CAGAGAGGCTATAGAGGGTTTTTCGCCT	656
Db	500 TGTCAACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCT	559
Qy	657 CAGTGCCCACTCTTCCTCTGACAGCAAAGATAACAC ACTCCACGTGGAAATTAACGGGTT	716
Db	560 TAGCGCCCACTGCTCCTGTGACAGCAGGGATAAACAC ACTGCAAGTGGACATCAACGGGTT	619
Qy	717 CAATTCTGGCCGCCGGGTGACCTGGCCACCATTCACGG CATGAACCGGCCCTTTCCTGCT	776
Db	620 CACTACCGGCCGCCGAGGTGACCTGGCCACCATTCATGG CATGAACCGGCCCTTTCCTGCT	679
Qy	777 CCTCATGGCCACCCCGCTGGAGAGGGCC CAGCCTG CACAGCTCCC GG CACCGCCGAG-	835
Db	680 TCTCATGGCCACCCCGCTGGAGAGGGCC CAGCATCTGCAAAGCTCCC GG CACCGCCGAGC	739
Qy	836 CCCTGGATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGA	895
Db	740 CCCTGGACACCAACTATTGCTTCAGCTCCCC-----GGAGA	775
Qy	896 AGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCG GAAGGACCTGGGCTGGAAGTGGA	955
Db	776 AGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCG CAGGGACCTCGGCTGGAAGTGGA	835
Qy	956 TTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGA	1015
Db	836 TCCACGAGCCCCAAGGCTACCATGCCACTTCTGCCTCGGGCCCTGCCCTACATTGGA	895
Qy	1016 GCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGT	1075
Db	896 GCCTGGACACGCGTACCAGCAAGTCCCTGGCCCTGTACACCCGCCATAACCGGGCGCCCT	955
Qy	1076 CGGCGGCGCCGTGCTGC- -GTGCCGAGGCGCTGGAGCCACTGCCCAT--CGTGTACTAC	1131
Db	956 CGGCGGCGCCGGTTTGCTGGGCCCCAGGCCCTTGACCCCTGCCCCCTTCGGGGTACTAC	1015

Qy 1132 GTGGGCCCGC--AAGCCCAAGGTGGAGCAGCTGTCCAACAT 1169  
 |||| ||| | ||||| ||| || ||||| ||||  
 Db 1016 CTGGGGCGCCAAGCCCCAAGTGGGAACACCTGTCCCACAT 1055

RESULT 2

BX349319

LOCUS BX349319 900 bp mRNA linear EST 05-MAY-2003

DEFINITION BX349319 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED  
 Homo sapiens cDNA clone CS0DL010YL07 5-PRIME, mRNA sequence.

ACCESSION BX349319

VERSION BX349319.1 GI:30379410

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 900)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 9160.r For  
 more information about this cluster, see

<http://www.genoscope.cns.fr/>

[cgi-bin/cluster.cgi?seq=CS0BAG059ZD04\\_CS05596\\_1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAG059ZD04_CS05596_1&cluster=9160.r).

Contact : Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0BAG059ZD04\_CS05596\_1.

FEATURES Location/Qualifiers

source

1. .900

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DL010YL07"

/cell\_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"

/cell\_line="RAMOS CELL LINE"

/clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT  
 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 179 a 292 c 270 g 155 t 4 others

ORIGIN

Query Match 58.6%; Score 702; DB 13; Length 900;

Best Local Similarity 88.0%; Pred. No. 3.5e-148;

Matches 762; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60  
 ||| |||| || ||||||||| ||||||||| ||||||||| |||

[illegible]

BX355682/c

DEFINITION BX355682 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI002YJ05 3-PRIME, mRNA sequence.

VERSION BX355682.1 GI:30371987

SOURCE Homo sapiens (human)

REFERENCE 1 (bases 1 to 1201)

TITLE Full-length cDNA libraries and normalization

COMMENT            Contact: Genoscope

Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 9160.r For  
more information about this cluster, see  
[http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DI002CE03NP1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DI002CE03NP1&cluster=9160.r). Contact :  
Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DI002CE03NP1.

## source

1. .1201

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9606"
```

```
/clone="CS0DI002YJ05"
```

/tissue type="PLACENTA COT 25-NORMALIZED"

```
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
```

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT	198 a	326 c	386 q	253 t	38 others
------------	-------	-------	-------	-------	-----------

## ORIGIN

Query Match 57.8%; Score 692.2; DB 13; Length 1201;

Best Local Similarity 86.7%; Pred. No. 6.3e-146:

Matches 824; Conservative 4; Mismatches 93; Indels 29; Gaps 5;

Qy 250 ACCCGCGACCGGGTAGCCGGGGAAAGTGTCTGAACCGGAGCCCGAGCCAGAG-GCGGACTA 308

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	5
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	---

1018 TATCCCGGACCCCCCTGGCCCGG AGAATTCAGATCCCGAGTCCCAGACTTGAGCCTCGACTA 969

Oy 309 CTACGCCAAGGAGGTCA CCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAA 368

Db 988 CTACGCCAAGGAGGTMACCCGCGTGCTAATGGTGGAAACCCACAACGAAATCTATGACAA 929

Qy	369	ATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGC	428
Db	928	GTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGC	869
Qy	429	GGTGCCGGAACCTGTATTGCTCT-CTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAA	487
Db	868	GGTACCTGAACCCGTGTTGCTCTCCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAA	809
Qy	488	AAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCA	547
Db	808	AAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCA	749
Qy	548	GCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGATGTCACCGGAG	607
Db	748	GCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTCACCGGAG	689
Qy	608	TTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACT	667
Db	688	TTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGA-GGCTTTTCGCCTTAGCGCCCACT	630
Qy	668	CTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCC	727
Db	629	GCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCC	570
Qy	728	GCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGGCCA	787
Db	569	GCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTCATGGCCA	510
Qy	788	CCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCA	847
Db	509	CCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGACACCA	450
Qy	848	ACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAACTGCTGCG	907
Db	449	ACTATTGCTTCAGCTCCAC-----GGAGAAGNACTGCTGCG	414
Qy	908	TGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCA	967
Db	413	TGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCA	354
Qy	968	AGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTC	1027
Db	353	AGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCCTACATTTGGAGCCTGGACACGC	294
Qy	1028	AGTACAGCAAGGTCTCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGT	1087
Db	293	AGTACAGCAAGGTCTCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGT	234
Qy	1088	GCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGCGCAAGCCCA	1147
Db	233	GCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCGCGCAAGCCCA	174
Qy	1148	AGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	1197
Db	173	AGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA	124

# RESULT 4

BX335351/c

LOCUS BX335351 983 bp mRNA linear EST 01-MAY-2003

DEFINITION BX335351 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI013YH16 3-PRIME, mRNA sequence.

ACCESSION BX335351

VERSION BX335351.1 GI:30308367

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 983)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 9160.r For more information about this cluster, see

[http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DI013DD08NP1](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DI013DD08NP1&cluster=9160.r)

&cluster=9160.r. Contact : Feng Liang Email : fliang@lifetech.com

URL : <http://fulllength.invitrogen.com/> InVitroGen Corporation 1600 Paraday Avenue Genoscope sequence ID : CS0DI013DD08NP1.

## FEATURES

source

Location/Qualifiers

1..983

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI013YH16"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 170 a 280 c 323 g 203 t 7 others

## ORIGIN

Query Match 57.3%; Score 685.8; DB 13; Length 983;

Best Local Similarity 86.6%; Pred. No. 1.6e-144;

Matches 805; Conservative 5; Mismatches 93; Indels 27; Gaps 4;

Qy 268 GGGGAAAGTGTGCGAACCGGAGCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCAACC 327

||||| ||||| ||||||||||||||||||| ||||| |||||||||||||||||||

Db 983 GGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGCCGACTACTACGCCAAGGAGGTCAACC 924

Qy 328 CGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAATTCAAGGGCACCCCCAC 387

||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 923 CGCGTGCTAATGGTGGARACCCACAACGAAATCTATGACAAGTTCAAGCAGAGTACACAC 864

Qy 388 AGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTG 447

||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 863 AGCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGTTG 804

Qy	448	CTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAAGTGGAGCAGCACGTGGAG	507
Db	803	CTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAAGTGGAGCAGCACGTGGAG	744
Qy	508	CTATACCAGAAATACAGCAATGATTCCCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCC	567
Db	743	CTGTACCAGAAATACAGCAACAATYCCTGGCGATACCTCAGCAACCGGCTGCTGGCACCC	684
Qy	568	AGTGA CTACCGGAGTGGCTGTCTTTGATGTACCGGAGTTGTGCGGCAGTGGCTGACC	627
Db	683	AGCGACTCGCCAGAGTGGTTATCTTTTGTATGTACCGGAGTTGTGCGGCAGTGGTTGAGC	624
Qy	628	CGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTCTTCTCTGACAGCAAAGAT	687
Db	623	CGTGGAGGGGAAATTGA-GGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGAT	565
Qy	688	AACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACC	747
Db	564	AACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCACC	505
Qy	748	ATTCACGGCATGAACCGGCCCTTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAG	807
Db	504	ATTCATGGCATGAACCGGCCCTTCTGCTTCTCATGGCCACCCCGCTGGAGAGGGCCAG	445
Qy	808	CACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACTACTGCTTCAGCTCCACG	867
Db	444	CATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGACACCAACTATTGCTTCAGCTCCAC-	386
Qy	868	GACTACAAGGATGACGACGACAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGAC	927
Db	385	-----GGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGAC	349
Qy	928	TTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTT	987
Db	348	TTCCGGAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTT	289
Qy	988	TGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCT	1047
Db	288	TGCCTCGGGCCCTGCCCCTACATTTGGAGCCTGGACACGCAGTAMMMCAAGGTCTGGCC	229
Qy	1048	CTGTACAACCAGCACAACCCGGGCGCGTTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTG	1107
Db	228	MTGTACAACCAGCATAACCC-GGCGCCTCGGCGGCGCCGTNGT-CGTGCCGAGGCGCTG	171
Qy	1108	GAGCCACTGCCCATCGTGTACTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAAC	1167
Db	170	GAGCCGCTGCCCATCGTGTACTACGTGGGCTGCAAGCCCAAGGTGGAGCAGCTGTCCAAC	111
Qy	1168	ATGATCGTGCGTTCTGCAAGTGCAGCTGA	1197
Db	110	ATGATCGTGCGTCTCTGCAAGTGCAGCTGA	81



DEFINITION BX324511 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
cDNA clone CS0DC024YD20 3-PRIME, mRNA sequence.

ACCESSION BX324511

VERSION BX324511.1 GI:30332381

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1041)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 9160.r For  
more information about this cluster, see  
[http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0AC024DB10NP2&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0AC024DB10NP2&cluster=9160.r). Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0AC024DB10NP2.

FEATURES Location/Qualifiers

source 1. .1041  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC024YD20"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 176 a 283 c 346 g 219 t 17 others

ORIGIN

Query Match 52.6%; Score 629.6; DB 13; Length 1041;  
Best Local Similarity 84.7%; Pred. No. 8.3e-132;  
Matches 788; Conservative 8; Mismatches 102; Indels 32; Gaps 7;

Qy 269 GGGAAAGTGTCTGAACCGGAGCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTACCCC 328  
:|| | :| ||||| ||||| || :|: | : || | |: |||||

Db 1022 SGGGAGAKGSAGAACCGGAGCCCGACCCCTRASCCR---ACTAYACCCCAAGRAGTCACCC 966

Qy 329 GCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAATTCAAGGGCACCCCCCACA 388  
||| |:||| |:| |||| | |||| | ||||| || ||||| | | ||||

Db 965 GCGTGYTAATGKT-GAAACCCACAACGAAATCTATGACAAGTTCAAGCAGAGTACACACA 907

Qy 389 GCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGC 448  
|| ||||| || | ||||| || ||||| ||||| ||||| || ||||| || |||||

Db 906 GCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGTTGC 847

Qy 449 TCTCTC-GGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAG 507

Db	846	TCTCCCSGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAG	787
Qy	508	CTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCC	567
Db	786	CTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCAGCAACCGGCTGCTGGCACCC	727
Qy	568	AGTGACTCACCGGAGTGGCTGTCTTTGATGTACCGGAGTTGTGCGGCAGTGGCTGACC	627
Db	726	AGCGACTCGCCAGAGTGGTTATCTTTGATGTACCGGAGTTGTGCGGCAGTGGTTGAGC	667
Qy	628	CGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTCTTCCTCTGACAGCAAAGAT	687
Db	666	CGTGGAGGGGAAATTGA-GGCTTTGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGAT	608
Qy	688	AACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACC	747
Db	607	AACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCACC	548
Qy	748	ATTCACGGCATGAACCGGCCCTTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAG	807
Db	547	ATTCATGGCATGAACCGGCCCTTCTGCTTCTCATGGCCACCCCGCTGGAGAGGGCCAG	488
Qy	808	CACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACTACTGCTTCAGCTCCACG	867
Db	487	CATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGACACCAACTATTGCTTCAGCTCCAC-	429
Qy	868	GACTACAAGGATGACGACGACAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGAC	927
Db	428	-----GGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGAC	392
Qy	928	TTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTC	987
Db	391	TTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTC	332
Qy	988	TGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCT	1047
Db	331	TGCCTCGGGCCCTGCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCC	272
Qy	1048	CTGTACAACCAGCACAAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTG	1107
Db	271	CTGTACAACCAGCATAACCC-GGCGCCTCGGCGGCGCCGTG-TGCGTGCCGCAGGCGCTG	214
Qy	1108	GAGCCACTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAAC	1167
Db	213	GAGCCGCTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAAC	154
Qy	1168	ATGATCGTGCGTTTCTGCAAGTGCAGCTGA	1197
Db	153	ATGATCGTGCGCTCCTGCAAGTGCAGCTGA	124

```

ACCESSION      BM555996
VERSION        BM555996.1  GI:18796907
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1  (bases 1 to 1043)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-r@mail.nih.gov
                Tissue Procurement: ATCC
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLAM12261 row: i column: 16
                High quality sequence stop: 702.

```

BASE COUNT	202 a	329 c	345 g	154 t	13 others
ORIGIN					

[illegible]

```

Db      181 CCAGCGACTCGCCAGAGTGGTTATCTTTTGTATGTCACCGGAGTTGTGCGGCAGTGGTTGA 240
Qy      626 CCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTCTTCCTCTGACAGCAAAG 685
      ||| ||| || || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      241 GCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGG 300
Qy      686 ATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCA 745
      ||||| ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db      301 ATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCA 360
Qy      746 CCATTACGGCATGAACCGGCCCTTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCC 805
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      361 CCATTGATGGCATGAACCGGCCCTTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCC 420
Qy      806 AGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACTACTGCTTCAGCTCCA 865
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      421 AGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGACACCAACTATTGCTTCAGCTCCA 480
Qy      866 CGGACTACAAGGATGACGACGACAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTG 925
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      481 C-----GGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTG 516
Qy      926 ACTTCCGGAAGGACCTGGGCTGGAAGTGGATTGATGAACCCAAGGGCTACCATGCCAATT 985
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      517 ACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACT 576
Qy      986 TCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGG 1045
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      577 TCTGCCTCGGGCCCTGCCCCTACATTGAGCCTGGACACGCAGTACAGCAAGGTCTCTGG 636
Qy      1046 CTCTGTACAACCAGCACAACCCGGGCGCGTGGCGGCGCCGTGCTGCGTGCCGAGGCGC 1105
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      637 CCCTGTACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGC 696
Qy      1106 TGGAGCCACTGCCCATCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCA 1165
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      697 TGGAGCCGCTGCCCATCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCA 756
Qy      1166 ACATGATCGTGCCTTCCTGCA 1186
      ||||| ||||| ||||| |||||
Db      757 ACATGATCGTGCCTTCCTGCA 777

```

# RESULT 7

BX383773/c

LOCUS BX383773 950 bp mRNA linear EST 08-MAY-2003

DEFINITION BX383773 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens  
cDNA clone CS0DK001YA15 3-PRIME, mRNA sequence.

ACCESSION BX383773

VERSION BX383773.1 GI:30457168

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 950)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.



Qy	703	GAAATTAACGGGTTC AATTCTGGCCGCCGCGGGTGACCTGGCCACCATTACGGCATGAAC	762
Db	584	GACATCAACGGGTTTACTACCGGCCGCCGAGGTGACCTGGCCACCATT CATGGCATGAAC	525
Qy	763	CGGCCCTTCTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCC	822
Db	524	CGGCCTTTCTCTGCTTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCC	465
Qy	823	CGGCACCGCCGAGCCCTGGATAACCACTACTGCTTCAGCTCCACGGACTACAAGGATGAC	882
		:	
Db	464	CGGCACCGCCGAGCMCTGGACACCAACTATTGCTTCAGCTCCAC-----	421
Qy	883	GACGACAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTG	942
		:	
Db	420	-----GGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTKACTTCCGCAAGGACCTC	369
Qy	943	GGCTGGAAGTGGATT CATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGT	1002
Db	368	GGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCGACTTCTGCCTCGGGCCCTGC	309
Qy	1003	CCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCAC	1062
Db	308	CCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCGCCTGTACAACCAGCA	249
Qy	1063	AACCCGGGCGCGTTCGGCGGCGCCGTGCTGCGTGCCG CAGGCGCTGGAGCCACTGCCCATC	1122
Db	248	T AACCCGGGCGCCTCGGCGGCGCCGTG-TGCGTGCCG CAGGCGCTGGAGCCGCTGCCCATC	190
Qy	1123	GTGTACTACGTGGGCCGCAAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCC	1182
		:	
Db	189	GTGTACTACGTGGGCKKAAGCCCCAAGGTGGAGCAGCTGTCCNACATGATCGTGCGCTCC	130
Qy	1183	TGCAAGTG CAGCTGA	1197
Db	129	TGCAAGTG CAGCTGA	115

Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 9160.r For  
 more information about this cluster, see  
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAK028AB08NM1&cluster=9160.r)  
[cgi-bin/cluster.cgi?seq=CS0BAK028AB08NM1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAK028AB08NM1&cluster=9160.r). Contact :  
 Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0BAK028AB08NM1.

FEATURES                      Location/Qualifiers  
     source                    1. .888  
                               /organism="Homo sapiens"  
                               /mol\_type="mRNA"  
                               /db\_xref="taxon:9606"  
                               /clone="CS0DE014YE16"  
                               /tissue\_type="PLACENTA"  
                               /clone\_lib="Homo sapiens PLACENTA"  
                               /note="Vector: pCMVSPORT\_6; 1st strand cDNA was primed  
                               with a NotI-oligo(dT) primer. Five prime end enriched,  
                               double-strand cDNA was digested with Not I and cloned into  
                               the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
                               Library was not normalized."

BASE COUNT            156 a       253 c       302 g       176 t       1 others

ORIGIN

Query Match                      49.2%;    Score 589;    DB 13;    Length 888;  
 Best Local Similarity       87.1%;    Pred. No. 1.2e-122;  
 Matches 679;    Conservative    0;    Mismatches    76;    Indels    25;    Gaps    2;

```

Qy      418 CTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGG 477
      ||||| ||||| || ||||| | ||| | ||||| ||||| ||||| |||||
Db      888 CTCCAAAAGCGGTACCTGAACCTTGTGTCTNCCCGGGCAGAGCTGCGTCTGCTGA-G 830

Qy      478 CTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGG 537
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      829 CTCAAGTTTAAAATGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGG 770

Qy      538 CGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGAT 597
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      769 CGATACCTCAGCAACCGGCTGCTGGCACACAGCGACTCGCCAGAGTGGTTATCTTTTGAT 710

Qy      598 GTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCCTC 657
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      709 GTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTT 650

Qy      658 AGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGGTTC 717
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      649 AGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGGTTC 590

Qy      718 AATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTC 777
      | | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      589 ACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTTCCTGCTT 530

Qy      778 CTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC 837
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      529 CTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCC 470

```

```

Qy      838 CTGGATACCAACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAG 897
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      469 CTGGACACCAACTATTGCTTCAGCTCCACG-----GAGAAG 434

Qy      898 AACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATT 957
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      433 AACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATC 374

Qy      958 CATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGC 1017
        || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      373 CACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGC 314

Qy      1018 CTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCG 1077
        || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      313 CTGGACACGCGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCG 254

Qy      1078 GCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGC 1137
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      253 GCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGC 194

Qy      1138 CGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAAGTGCAGCTGA 1197
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      193 CGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTCTCTGCAAGTGCAGCTGA 134

```

# RESULT 9

AL530081

LOCUS AL530081 859 bp mRNA linear EST 23-MAY-2003

DEFINITION AL530081 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens cDNA clone CS0DD009YM06 5-PRIME, mRNA sequence.

ACCESSION AL530081

VERSION AL530081.2 GI:31067916

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 859)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT On Feb 13, 2001 this sequence version replaced gi:12793574.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9160.r For more information about this cluster, see

<http://www.genoscope.cns.fr/>

[cgi-bin/cluster.cgi?seq=CS0DD009BG03QP1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DD009BG03QP1&cluster=9160.r). Contact :

Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DD009BG03QP1.

FEATURES Location/Qualifiers

source 1..859

/organism="Homo sapiens"



/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DD009YM06"  
/tissue\_type="NEUROBLASTOMA COT 50-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT        179 a        272 c        257 g        148 t        3 others  
ORIGIN

Query Match                    48.2%;    Score 577.2;    DB 9;    Length 859;  
Best Local Similarity    87.3%;    Pred. No. 5.3e-120;  
Matches 641;    Conservative    2;    Mismatches    90;    Indels    1;    Gaps    1;

```
Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| |||| || ||||||||| ||| ||||||||| ||||||||| ||| |||
Db     126 ATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 185

Qy      61 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||||||||| ||||||||| ||||| ||||||||| ||||||||| |||
Db     186 CTGACGCCTGGCCGGCCGGCCCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 245

Qy     121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||
Db     246 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 305

Qy     181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||||||||| ||||||||| ||||||||| ||||||||| ||||| ||| |||
Db     306 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCCAGGCCGTGCTCGCCCTG 365

Qy     241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG 300
      ||||||||| ||||||||| ||||||||| ||||| ||||||||| ||||||||| |||
Db     366 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 425

Qy     301 GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      || ||||||||| ||||||||| ||||||||| ||||||||| ||||| |||||
Db     426 GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC 485

Qy     361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420
      ||||| || ||||||| || ||||||| ||||||||| ||||||||| || |||||||
Db     486 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 545

Qy     421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
      || ||||||||| || ||||||||| ||||||||| ||||||||| |||||||||
Db     546 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 605

Qy     481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC 540
      ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||
Db     606 AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA 665

Qy     541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTG-GCTGTCCTTTGATGT 599
      ||||||||| ||||||||| ||||| ||||| || ||||| || || |||||
Db     666 TACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGTGTTATCTTTTGTATGT 725

Qy     600 CACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAG 659
```

Db 726 CACCGGAGTTGTGCGGCAGTKGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTCGCCTTAG 785  
 Qy 660 TGCCCACTCTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAA 719  
 Db 786 CGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGRACATCAACGGGTTTAC 845  
 Qy 720 TTCTGGCCGCCGGG 733  
 Db 846 TACCGGCCGCCGAG 859

RESULT 10

BE312000

LOCUS	BE312000	713 bp	mRNA	linear	EST 26-OCT-2000
-------	----------	--------	------	--------	-----------------

DEFINITION 601154768F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:3510592 5', mRNA sequence.

ACCESSION BE312000

VERSION BE312000.1 GI:9130128

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 713)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL      Unpublished

COMMENT      Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)

Plate: LLCM195 row: h column: 17

High quality sequence start: 2

High quality sequence stop: 713.

FEATURES	Location/Qualifiers
----------	---------------------

source

1. .713

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9606"
```

```
/clone="IMAGE:3510592"
```

```
/tissue type="neuroblastoma"
```

```
/lab_host="DH10B (phage-resistant)"
```

```
/clone_lib="NIH MGC 19"
```

```
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
```

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University o

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II  
Note: this is a NIH MGC Library."

BASE COUNT	157 a	226 c	211 g	119 t
------------	-------	-------	-------	-------

ORIGIN

Query Match 47.7%; Score 570.6; DB 10; Length 713;  
 Best Local Similarity 87.5%; Pred. No. 1.6e-118;  
 Matches 624; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```

Qy      125 AGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCC 184
          |||||
Db       1  AGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCC 60

Qy      185 CCCCAGGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACA 244
          |||||
Db       61 CCCCAGGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTGTACA 120

Qy      245 ACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGCCAGAGGCGG 304
          |||||
Db      121 ACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGCCG 180

Qy      305 ACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATG 364
          |||||
Db      181 ACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATCTATG 240

Qy      365 ATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGG 424
          |||||
Db      241 ACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAG 300

Qy      425 AAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGT 484
          |||||
Db      301 AAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGT 360

Qy      485 TAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGCTACC 544
          |||||
Db      361 TAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGATAACC 420

Qy      545 TCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTCAACG 604
          |||||
Db      421 TCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTCAACG 480

Qy      605 GAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTGCGCTCAGTGCCC 664
          |||||
Db      481 GAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTGCGCTTAGCGCCC 540

Qy      665 ACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTG 724
          |||||
Db      541 ACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCG 600

Qy      725 GCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGCTCCTCATGG 784
          |||||
Db      601 GCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGCTTCTCATGG 660

Qy      785 CCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC 837
          |||||
Db      661 CCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCGAGCC 713
  
```

LOCUS BE260971 717 bp mRNA linear EST 26-OCT-2000  
 DEFINITION 601153715F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:3509931 5',  
 mRNA sequence.  
 ACCESSION BE260971  
 VERSION BE260971.1 GI:9132709  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 717)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: LLCM193 row: m column: 04  
 High quality sequence stop: 713.  
 FEATURES  
 source Location/Qualifiers  
 1. .717  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3509931"  
 /tissue\_type="neuroblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_19"  
 /note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."  
 BASE COUNT 158 a 226 c 213 g 120 t  
 ORIGIN

Query Match 47.0%; Score 562.6; DB 10; Length 717;  
 Best Local Similarity 87.4%; Pred. No. 1e-116;  
 Matches 627; Conservative 0; Mismatches 89; Indels 1; Gaps 1;

Qy 126 GCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCC 185  
 |||  
 Db 1 GCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCCC 60  
 Qy 186 CCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGGCTCTTTACAA 245  
 |||  
 Db 61 CCCGAGCCAGG-GGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCGCCCTGTACAA 119  
 Qy 246 CAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAGGCGGA 305

Source	Position	Sequence	Position
Db	120	CAGCACCCGCGACCCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGCCGA	179
Qy	306	CTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGA	365
Db	180	CTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATCTATGA	239
Qy	366	TAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGA	425
Db	240	CAAGTTC AAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGA	299
Qy	426	AGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTT	485
Db	300	AGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTT	359
Qy	486	AAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCT	545
Db	360	AAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGATACCT	419
Qy	546	CAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTACCCGG	605
Db	420	CAGCAACCGGCTGCTGGCAACCAGCGACTCGCCAGAGTGGTTATCTTTGATGTACCCGG	479
Qy	606	AGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCCTCAGTGCCCA	665
Db	480	AGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCTTAGCGCCCA	539
Qy	666	CTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGG	725
Db	540	CTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGG	599
Qy	726	CCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGGC	785
Db	600	CCGCCGAGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTTCTCATGGC	659
Qy	786	CACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGA	842
Db	660	CACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCGAGCCCTGGGA	716

RESULT 12

BQ675698

LOCUS	BQ675698	902 bp	mRNA	linear	EST 15-JUL-2002
-------	----------	--------	------	--------	-----------------

DEFINITION AGENCOURT\_8036532 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:6211917 5', mRNA sequence.

ACCESSION B0675698

VERSION BO675698.1 GI:21786532

KEYWORDS EST.

SOURCE	Homo sapiens (human)
--------	----------------------

ORGANISM	Homo sapiens
----------	--------------

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 902)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LCM2375 row: c column: 22  
High quality sequence stop: 599.

FEATURES  
source Location/Qualifiers  
1. .902  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6211917"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_102"  
/note="Organ: salivary gland; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGACGAG(G). Library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZAP-cDNA  
synthesis kit (Stratagene) and Superscript II RT (Life  
Technologies). Note: this is a NIH\_MGC Library."  
BASE COUNT 194 a 292 c 255 g 161 t  
ORIGIN

Query Match 46.9%; Score 561.6; DB 13; Length 902;  
Best Local Similarity 87.4%; Pred. No. 1.8e-116;  
Matches 615; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy	146	TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC	205
Db	1	TCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCCCCCGAGCCAGGGGGAGGTGC	60
Qy	206	CGCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG	265
Db	61	CGCCCGGCCCCGCTGCCCCGAGGCCGTGCTCGCCCTGTACAACAGCACCCGCGACCGGGTGG	120
Qy	266	CCGGGGAAAGTGTCGAACCGGAGCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCA	325
Db	121	CCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGCCGACTACTACGCCAAGGAGGTCA	180
Qy	326	CCCGCGTGCTAATGGTGGAAGCGGCAACCAAATCTATGATAAATTCAAGGGCACCCCCC	385
Db	181	CCCGCGTGCTAATGGTGGAACCCACAACGAAATCTATGACAAGTTCAAGCAGAGTACAC	240
Qy	386	ACAGCTTATATATGCTGTTCAACACGTCCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTAT	445
Db	241	ACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGT	300
Qy	446	TGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG	505
Db	301	TGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG	360

Qy	506	AGCTATAACAGAAATACAGCAATGATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCC	565
Db	361	AGCTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCAGCAACCGGCTGCTGGCAC	420
Qy	566	CCAGTGACTACCGGAGTGGCTGTCTTTGATGTACCGGAGTTGTGCGGCAGTGGCTGA	625
Db	421	CCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTACCGGAGTTGTGCGGCAGTGGTTGA	480
Qy	626	CCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTCTTCCTCTGACAGCAAAG	685
Db	481	GCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGG	540
Qy	686	ATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCA	745
Db	541	ATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCA	600
Qy	746	CCATTCACGGCATGAACCGGCCCTTCCTGCTCCTCATGGCCACCCGCTGGAGAGGGCCC	805
Db	601	CCATTCATGGCATGAACCGGCCCTTCCTGCTTCTCATGGCCACCCGCTGGAGAGGGCCC	660
Qy	806	AGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAAC	849
Db	661	AGCATCTGCCAAGCTCCCGGCACCGCCGAGCCCTGGGACACCAC	704

RESULT 13

BI 818841

LOCUS	BI818841	925 bp	mRNA	linear	EST 04-OCT-2001
-------	----------	--------	------	--------	-----------------

DEFINITION 603037307F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5178433 5', mRNA sequence.

ACCESSION BI818841

VERSION BI818841.1 GI:15930391

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 925)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL      Unpublished

COMMENT            Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11444 row: n column: 02

High quality sequence stop: 874.

FEATURES	Location/Qualifiers
----------	---------------------

```
source      1.  .925
```

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```

/db_xref="taxon:9606"
/clone="IMAGE:5178433"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

```

```

BASE COUNT      170 a      316 c      278 g      161 t
ORIGIN

```

```

Query Match          44.6%;  Score 534.2;  DB 12;  Length 925;
Best Local Similarity 87.6%;  Pred. No. 2.8e-110;
Matches 607;  Conservative 0;  Mismatches 83;  Indels 3;  Gaps 2;

```

```

Qy      1 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60
      ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     151 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 210

Qy      61 CTGACGCCTGGCCGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     211 CTGACGCCTGGCCGGCCGGCCGGCCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG 270

Qy     121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     271 GTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 330

Qy     181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     331 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG 390

Qy     241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     391 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG 450

Qy     301 GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC 360
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     451 GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC 510

Qy     361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420
      ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     511 TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 570

Qy     421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480
      || ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db     571 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC 630

Qy     481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC 540
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     631 AAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA 690

```



```

Qy      541 TACCTCAGCAACCGGCTGC-TGGCCCCCAGTGACTCACCGGAGTGGCTGTCTTTTGATGT 599
          |||
Db      691 TACCTCAGCAACCGGCTGCTTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTTGATGT 750
          |||
Qy      600 CACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAG--GAGGCTATAGAGGGTTTTCGCCTC 657
          |||
Db      751 CACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGACGGGGAAATTGAGGGCTTTTCGCTT 810
          |||
Qy      658 AGTGCCCACTCTTCCTCTGACAGCAAAGATAAC 690
          ||
Db      811 AGCGCCCACTGCTCCTGTGACAGCAGGGATAACC 843
          ||

```

RESULT 14

B0054305

LOCUS	B0054305	1093 bp	mRNA	linear	EST 29-MAR-2002
-------	----------	---------	------	--------	-----------------

DEFINITION AGENCOURT\_6830409 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936397  
5', mRNA sequence.

ACCESSION      B0054305

VERSION BO054305.1 GI:19813645

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1093)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT            Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Dr. Daniel McVicar, DBS/NCI

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2125 row: k column: 22

High quality sequence stop: 605.

FEATURES	Location/Qualifiers
----------	---------------------

source

1. .1093

/organis

```
/mol type="mRNA"
```

```
/db xref="taxon:
```

/clone="IMAGE:5936397

```
/tissue_type="natural killer cells, cell line"
```

```
/lab host="DH10B (phage-resistant)"
```

```
/clone lib="NIH MGC 106"
```

```
/note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
```

EcoRI; cDNA made by oligo-dT priming. Directionally clone into EcoRI/XhoI sites using the following 5' adaptor:

GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California.

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a

NIH MGC Library."



AL530080/c  
 LOCUS AL530080 841 bp mRNA linear EST 23-MAY-2003  
 DEFINITION AL530080 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens  
 cDNA clone CS0DD009YM06 3-PRIME, mRNA sequence.  
 ACCESSION AL530080  
 VERSION AL530080.2 GI:31067915  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 841)  
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT On Feb 13, 2001 this sequence version replaced gi:12793573.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 9160.r For more information about this cluster, see  
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DD009BG03NP1&cluster=9160.r)  
[cgi-bin/cluster.cgi?seq=CS0DD009BG03NP1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DD009BG03NP1&cluster=9160.r). Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DD009BG03NP1.  
 FEATURES Location/Qualifiers  
 source 1. .841  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DD009YM06"  
 /tissue\_type="NEUROBLASTOMA COT 50-NORMALIZED"  
 /clone\_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 150 a 240 c 283 g 153 t 15 others  
 ORIGIN

Query Match 43.5%; Score 520.2; DB 9; Length 841;  
 Best Local Similarity 84.0%; Pred. No. 4e-107;  
 Matches 647; Conservative 7; Mismatches 86; Indels 30; Gaps 5;

Qy 428 CGGTGCCGGAACCTGTATTGCTCT-CTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTA 486  
 |||| | | :|| | |||:| | | ||||| ||||| |||||  
 Db 841 CGGTACCTGMACCCTTGTTGYTCTCCCCGGGCAGAGCTGCTTCTGCTAAGGCTCAAGTTA 782  
 Qy 487 AAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGCTACCTC 546  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 781 AAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATWCCTGGCGGATACCTC 722  
 Qy 547 AGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCCTTTGATGTACCGGA 606  
 ||||| :|| ||||| ||||| :||| || ||||| | || ||||| ||||| :|

Db 721 AGCAACCSGCSGCTGGCACCCAGCSACTCGCCAGAGTGGTTTTCTTTTGTGTGTCACCGKA 662

Qy 607 GTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGTGCCAC 666  
 |||||

Db 661 GTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTAAGG--TTTCGCCTTAGCGCCNAC 604

Qy 667 TCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCATTCTGGC 726  
 |||||

Db 603 GGCTCCTGTAAACAGCAGGGATAACACACTGCAAGTGAACATCAACGGGTAACTACCGGC 544

Qy 727 CGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGGCC 786  
 |||||

Db 543 CGCCGAGGTGACCTGGCCACCATTCTTGGCATGAACCGGCCCTTCCTGCTTCTCATGGCC 484

Qy 787 ACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACC 846  
 |||||

Db 483 ACCCCGCTGGAGAGGGCCCAGCTTCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGACACC 424

Qy 847 AACTACTGCTTCAGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAACTGCTGC 906  
 |||||

Db 423 AACTATTGCTTCAGCTCCACGG-----AGAAGAACTGCTGC 388

Qy 907 GTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCC 966  
 |||||

Db 387 GTGCGGCAGCTGTACATTAACTTCCGCAAGGACCTCGGCTGGGAGTGGATCCACGAGCCC 328

Qy 967 AAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACT 1026  
 |||||:

Db 327 AAGGGCTWCCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTACATTTGGAGCCTGGACACG 268

Qy 1027 CAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCG 1086  
 |||

Db 267 CAGGACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGG--GCCTCGGCGGCGCCG 210

Qy 1087 TGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGGCAAGCCC 1146  
 |||

Db 209 TGC-GCGTGCCGAGGCGCTGGAGCCGCTGCCCTTCGTGTACTACGTGGGCGGCAAGCCC 151

Qy 1147 AAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTG 1196  
 |||||

Db 150 AAGGTGGGGCGGCTGTCCAACATGGTTCGTGCGCTCCTGCAAGGGCCGCTG 101

Search completed: October 28, 2003, 00:08:13  
 Job time : 2974.42 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

```
Run on:      October 27, 2003, 19:11:58 ; Search time 33.2026 Seconds
              (without alignments)
              1902.657 Million cell updates/sec
```

Title: US-10-017-372E-13  
Perfect score: 2114  
Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 398

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	2053	97.1	390	23	AAE13596		Porcine transformi
2	1931	91.3	390	22	AAM39186		Human polypeptide
3	1924	91.0	390	7	AAP61468		PreTGF-beta gene p
4	1924	91.0	390	11	AAR04034		Sequence of pre-TG
5	1924	91.0	390	11	AAR05258		Human pre-transfor
6	1924	91.0	390	12	AAR13813		Human pro-TGF-beta
7	1924	91.0	390	16	AAR73596		Human TGF-beta 1 p
8	1924	91.0	390	17	AAR90827		Pre-transforming g
9	1921	90.9	390	23	AAU77101		Human transforming
10	1921	90.9	390	23	AAE16943		Human transforming
11	1920.5	90.8	391	24	ABB82780		TGFB1 Arg25Pro pol
12	1919	90.8	390	13	AAR20124		Sequence of simian
13	1915	90.6	390	15	AAR46227		Human pre-TGF-beta
14	1913.5	90.5	391	16	AAR83054		Transforming growt
15	1913	90.5	390	19	AAW78785		Human pre-transfor
16	1912	90.4	390	22	AAB84601		Nucleotide sequenc
17	1909.5	90.3	391	9	AAP81362		Human transforming
18	1908.5	90.3	434	11	AAR03743		Monkey transformin
19	1905	90.1	390	24	ABB82781		TGFB1 Arg25Pro pol
20	1888	89.3	386	11	AAR05663		Simian Transformin
21	1882.5	89.0	387	11	AAR05664		Simian Transformin
22	1872	88.6	390	11	AAR05492		Chimeric simian TG
23	1868	88.4	390	13	AAR27522		TGF-beta 1/beta 2
24	1852.5	87.6	453	22	AAM40972		Human polypeptide
25	1848.5	87.4	389	13	AAR29657		TGF-beta 1. Homo
26	1763	83.4	390	11	AAR05749		Human TGF-Beta2 ex
27	1763	83.4	390	13	AAR20126		Sequence of hybrid
28	1757	83.1	390	11	AAR05665		Human Transforming
29	1755	83.0	390	11	AAR05666		Hybrid transformin
30	1747.5	82.7	391	10	AAP91900		Sequence encoded b
31	1719.5	81.3	389	16	AAR79921		Simian-human hybri
32	1300	61.5	278	15	AAR53090		Polypeptide cross-
33	1295	61.3	278	12	AAR12541		Latency associated
34	1262.5	59.7	458	23	ABG31507		LAP-mIFNB construc
35	1262.5	59.7	463	23	ABG31510		LAP-huIFNB constru
36	1173.5	55.5	290	22	ABG06792		Novel human diagno
37	1135	53.7	450	23	ABG31508		mIFNB-LAP construc
38	1118	52.9	448	23	ABG31509		huIFNB-LAP construc
39	944	44.7	227	22	ABG20234		Novel human diagno
40	896.5	42.4	236	22	ABG20233		Novel human diagno
41	874	41.3	382	21	AAB08338		Amino acid sequenc
42	874	41.3	382	23	AAU77105		Frog transforming
43	837	39.6	456	19	AAW78786		Pig transforming g
44	830.5	39.3	456	15	AAR46228		Pig TGF-beta-3. S
45	830	39.3	412	16	AAR73598		Human TGF-beta 3 p

#### ALIGNMENTS

RESULT 1

AAE13596

ID AAE13596 standard; Protein; 390 AA.

XX

AC AAE13596;

XX

DT 26-FEB-2002 (first entry)

XX

DE Porcine transforming growth factor beta 1 (TGF-beta1) mutant.

XX

KW Porcine; transforming growth factor beta 1; TGF-beta1; gene therapy;  
KW IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;  
KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;  
KW diabetes mellitus; sarcoidosis; psoriasis; dermatological; mutant;  
KW mutein.

XX

OS Sus scrofa.

XX

FH Key Location/Qualifiers

FT Misc-difference 223

FT /note= "Wild type Cys substituted with Ser"

FT Misc-difference 225

FT /note= "Wild type Cys substituted with Ser"

XX

PN WO200181404-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US12980.

XX

PR 20-APR-2000; 2000US-199014P.

XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Strober W, Nakamura K, Kitani A, Fuss IJ;

XX

DR WPI; 2002-026155/03.

DR N-PSDB; AAD22696.

XX

PT Composition for treating autoimmune diseases e.g. inflammatory bowel  
PT disease in humans, comprises vector containing transforming growth  
PT factor-beta under the control of inducible promoter -

XX

PS Example 1; Fig 1; 78pp; English.

XX

CC The invention relates to a composition containing a vector comprising a  
CC gene encoding a regulatory transcription factor under the control of a  
CC promoter encoding a transforming growth factor-beta (TGF-beta). The  
CC vector is useful for expressing TGF-beta, such as TGF-beta1, TGF-beta2  
CC or TGF-beta3, its variants or homologues, by transfecting a cell which  
CC is part of a host suspected of having an autoimmune disease, especially  
CC inflammatory bowel disease (IBD), under conditions such that the  
CC polypeptide encoded by the nucleic acid sequence in the vector is  
CC expressed. The vector is delivered using a delivery system. The delivery  
CC of the vector results in substantial elimination of symptoms of the  
CC autoimmune disease and increased production of IL-10 by the host. The  
CC composition is useful for treating various diseases with an autoimmune

CC component such as multiple sclerosis, rheumatoid arthritis, systemic  
CC lupus erythematosus, insulin-dependent diabetes mellitus, sarcoidosis  
CC and psoriasis, and also for assaying the expression of a gene in a cell.  
CC The vector is further useful for screening of the effect of test  
CC compounds on cytokine (e.g. TGF-beta) expression of transfected cells.  
CC The present sequence is porcine TGF-beta1 mutant.  
XX  
SQ Sequence 390 AA;

Query Match 97.1%; Score 2053; DB 23; Length 390;  
Best Local Similarity 98.0%; Pred. No. 1.5e-173;  
Matches 390; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA 60
      |||
Db      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA 60

Qy     61 SPFSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      |||
Db     61 SPFSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      |||
Db    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      |||
Db    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      |||
Db    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      |||
Db    293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy    361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      |||
Db    353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

## RESULT 2

AAM39186

ID AAM39186 standard; Protein; 390 AA.

XX

AC AAM39186;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 2331.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;



KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR N-PSDB; AAI58342.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Example 4; SEQ ID NO 2331; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 SQ Sequence 390 AA;

Query Match 91.3%; Score 1931; DB 22; Length 390;  
 Best Local Similarity 92.0%; Pred. No. 9.9e-163;  
 Matches 366; Conservative 10; Mismatches 14; Indels 8; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI R GQILSKLRLA 60

Db	1	MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
Db	61	SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN	300
Db	241	TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN	292
Qy	301	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	360
Db	293	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	352
Qy	361	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	398
Db	353	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 3

AAP61468

ID AAP61468 standard; Protein; 390 AA.

XX

AC AAP61468;

XX

DT 31-OCT-2002 (updated)

DT 28-OCT-1991 (first entry)

XX

DE PreTGF-beta gene product.

XX

KW Transforming growth factor beta; cancer; wound healing.

XX

OS Unidentified.

XX

FH Key Location/Qualifiers

FT Protein 279..390

XX

PN EP200341-A.

XX

PD 10-DEC-1986.

XX

PF 21-MAR-1986; 86EP-0302112.

XX

PR 22-MAR-1985; 85US-0715142.

PR 13-MAR-1987; 87US-0025423.

XX

PA (GETH ) GENENTECH INC.

XX

PI Derynck RMA;  
XX  
DR WPI; 1986-326875/50.  
DR N-PSDB; AAN60972.  
XX  
PT TGF-beta prodn. from transformed hosts - useful esp. for treating  
PT wounds (J6 2/9/86).  
XX  
PS Disclosure; Fig 1b; 26pp; English.  
XX  
CC The gene product is known to stimulate cell proliferation and  
CC inhibit anchorage-dependent growth of a variety of human cancer cell  
CC lines, it is esp. useful in treatment of burns and the promotion of  
CC surface and internal wound healing. TGF-beta may be expressed from a  
CC transformed CHO cell line.  
CC (Updated on 31-OCT-2002 to add missing OS field.)  
XX  
SQ Sequence 390 AA;

Query Match 91.0%; Score 1924; DB 7; Length 390;  
Best Local Similarity 91.7%; Pred. No. 4.1e-162;  
Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1;

Qy	1	MAPSGRLRLPLLLPLLWLLVLTGPPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Db	1	MPPSGRLRLPLLLPLLWLLVLTGPSPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Qy	61	SPPSQGDVPVPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMEVESGNQI	120
		:   :   :	
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMEVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:   :   :	
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNNSWR	180
Qy	181	YLSNRLLAPSDSPEWL SFDVTGVVRQWLTRREAI EGFRLSAHSSSDSKDNTLHVEINGFN	240
		:                 :       :	
Db	181	YLSNRLLAPSDSPEWL SFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN	300
		:	
Db	241	TGRRGDLATI HGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN	292
Qy	301	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	360
Db	293	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	352
Qy	361	APCCVPQALEPLPI VYYYVGRKPKVEQLSNMIVRSCKCS	398
Db	353	APCCVPQALEPLPI VYYYVGRKPKVEQLSNMIVRSCKCS	390

```

RESULT 4
AAR04034
ID      AAR04034 standard; protein; 390 AA.
XX
AC      AAR04034;

```

```

XX      25-MAR-2003   (updated)
XX      31-OCT-2002   (updated)
XX      31-MAY-1989   (first entry)
XX
DE      Sequence of pre-TGF-beta 1.
XX
KW      Transforming growth factor beta-3 (TGF beta 3); tumour cells; growth
KW      inhibition.
XX
OS      Homo sapiens.
XX
FH      Key           Location/Qualifiers
FT      CDS           348..500
XX
PN      WO8912101-A.
XX
PD      14-DEC-1989.
XX
PF      08-JUN-1988;   88WO-US01945.
XX
PR      08-JUN-1988;   88WO-US01945.
XX
PA      (GETH ) GENENTECH INC.
XX
PI      Dernyck RMA,   Goeddel DV;
XX
DR      WPI; 1990-007474/01.
DR      N-PSDB; AAQ02815.
XX
PT      Nucleotide sequence encoding transforming growth factor beta-3 used as a
PT      probe, or to produce TGF beta 3, for inhibiting growth of certain normal
PT      and neoplastic cells, eg A549.
XX
PS      Disclosure; Fig. 2; 61pp; English.
XX
CC      Sequence is an exon of transforming growth factor-beta 1 (pre-TGF-beta
CC      1) polypeptide and corresponds to AA's 288-338 of mature TGF-beta 1. The
CC      nucleic acid encoding second subtype of TGF-beta (TGF-beta 3) is useful
CC      as a probe or to produce TGF-beta 3 for inhibition of normal and
CC      neoplastic cell growth.
CC      (Updated on 31-OCT-2002 to add missing OS field.)
CC      (Updated on 25-MAR-2003 to correct PR field.)
CC      (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ      Sequence      390 AA;

Query Match          91.0%;   Score 1924;   DB 11;   Length 390;
Best Local Similarity 91.7%;   Pred. No. 4.1e-162;
Matches 365;   Conservative 10;   Mismatches 15;   Indels 8;   Gaps 1;

Qy      1  MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      |  |||||||||||||||||||
Db      1  MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
Qy      61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      |||||:|||||||||||||||||

```

Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:     :   :	
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
		:                         :         :	
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWL SRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMRNPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN	300
		:	
Db	241	TGRRGDLATIHGMRNPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN	292
Qy	301	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	360
Db	293	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	352
Qy	361	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	398
Db	353	APCCVPOALEPLPIVYYVGRKPKVEOLSNMIVRSCKCS	390

## RESULT 5

AAR05258

ID AAR05258 standard; protein; 390 AA.

XX

AC AAR05258;

XX

DT 25-MAR-2003 (updated)

DT 05-AUG-1990 (first entry)

XX

DE Human pre-transforming growth factor-beta-1 (pre-TGF-beta-1).

XX

KW Transforming growth factor-beta-1 (TGF-beta-1);

KW neoplastic cell line inhibition;

KW EGF-potentiated anchorage-independent growth.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	Peptide	1..278
----	---------	--------

FT	Protein	279..2011
----	---------	-----------

FT Domain 8..23

```
FT                               /note="hydrophobic domain"
```

FT Modified-site 82..84

FT /note="potential N-glycosylation site "

FT	Modified-site	136..138
----	---------------	----------

FT /note="as above"

FT	Modified-site	176..178
----	---------------	----------

FT /note="as above"

FT Cleavage-site 277..278

FT /note="proteolytic cleavage site"

XX

PN US4886747-A.

XX

PD 12-DEC-1989.

```
XX PF 13-MAR-1987; 87US-0025423.
XX
PR 13-MAR-1987; 87US-0025423.
PR 22-MAR-1985; 85US-0715142.
XX
PA (GETH ) GENENTECH INC.
XX
PI Derynck RMA, Goeddel DV;
XX
DR WPI; 1990-051338/07.
DR N-PSDB; AAQ93301.
XX
PT Nucleic acid encoding transforming growth factor-beta -
PT cloned into expression vectors for expression in eukaryotic host
PT cells for therapeutic use
XX
PS Disclosure; Fig 1b; 28pp; English.
XX
CC Pre-TGF-beta-1 is a polypeptide of 390 amino acids. Post-transitional
CC cleavage of the precursor gives rise to the mature TGF-beta monomer.
CC The sequence for human TGF-beta was determined by direct amino acid
CC sequence analysis and by deduction from the TGF-beta cDNA. It is
CC capable of inducing EGF-potentiased anchorage-independent growth of
CC target cell lines, and/or growth inhibition of neoplastic cell lines. It
CC can be used for treating wounds, eg burns or epidermal ulcers.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 390 AA;

Query Match          91.0%; Score 1924; DB 11; Length 390;
Best Local Similarity 91.7%; Pred. No. 4.1e-162;
Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1;

Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAGLSTCKTIDMELVKRKRIEAIRGQILSKRLRA 60
        | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTGPPPAGLSTCKTIDMELVKRKRIEAIRGQILSKRLRA 60

Qy     61 SPSPQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVL MVESGNQI 120
        ||||| : ||||| | | | | | | | | | | | | | | | | | | | | | : ||
Db     61 SPSPQGEVPPGGLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVL MVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
        ||||| : || : || | | | | | | | | | | | | | | | | | | | | : ||
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWL SFDVTGVVRQWL TRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
        ||||| | | | | | | | | | | | | | | | | | | | | | : || | | : ||||
Db    181 YLSNRLLAPSDSPEWL SFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240

Qy    241 SGRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
        : ||||| | | | | | | | | | | | | | | | | | | | | | |||
Db    241 TGRRGDLATI HGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
        ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352
```

Qy 361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398  
 |||||  
 Db 353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 6

AAR13813

ID AAR13813 standard; Protein; 390 AA.

XX

AC AAR13813;

XX

DT 20-NOV-1991 (first entry)

XX

DE Human pro-TGF-beta 1.

XX

KW Osteogenetic; tumoricidal.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..29

FT /note= "signal peptide"

FT Peptide 30..390

FT /note= "pro-TGF-beta 1"

FT Peptide 279..390

FT /note= "TGF-beta 1"

XX

PN JP03180192-A.

XX

PD 06-AUG-1991.

XX

PF 07-DEC-1989; 89JP-0318243.

XX

PR 07-DEC-1989; 89JP-0318243.

XX

PA (KIRI ) KIRIN BREWERY KK.

XX

DR WPI; 1991-271579/37.

DR N-PSDB; AAQ13392.

XX

PT Human pro-TGF-beta 1 prodn., for osteo-genetic activity - by

PT preparing DNA chain contg. base sequence coding for human

PT pre:pro-TGF-beta 1, forming expression vector etc.

XX

PS Claim 1; Fig 1; 16pp; Japanese.

XX

CC The amino acid sequence codes for human prepro-TGF-beta 1 which

CC can be produced by recombinant methods, it has osteogenetic and

CC tumoricidal activity.

XX

SQ Sequence 390 AA;

Query Match 91.0%; Score 1924; DB 12; Length 390;

Best Local Similarity 91.7%; Pred. No. 4.1e-162;

Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1;

[illegible]

RESULT 7

AAR73596

ID AAR73596 standard; Protein; 390 AA.

XX

AC AAR73596;

XX

DT 25-MAR-2003 (updated)

DT 20-DEC-1995 (first entry)

XX

DE Human TGF-beta 1 protein.

XX

KW Transforming growth factor-beta; Human TGF-beta protein; TGF-beta 1;

KW TGF-beta 3; osteogenic cell source; OCS; bone deficiency;

KW bone-inducing cofactor.

XX

OS Homo sapiens.

XX

PN US5409896-A.

XX

PD 25-APR-1995.

XX

PF 12-NOV-1993; 93US-0132405.

XX

PR 01-SEP-1989; 89US-0401906.

PR 12-NOV-1991; 91US-0790856.

PR 18-MAY-1993; 93US-0063841.

PR 12-NOV-1993; 93US-0132405.

XX



PA (GETH ) GENENTECH INC.  
 XX  
 PI Ammann AJ, Rudman CG;  
 XX  
 DR WPI; 1995-169610/22.  
 XX  
 PT Compsn. for treating skeletal tissue deficiency - comprising  
 PT transforming growth factor-beta and an osteogenic cell source in a  
 PT carrier  
 XX  
 PS Claim 3; Column 15-18; 19pp; English.  
 XX  
 CC This sequence represents human transforming growth factor-beta 1  
 CC (TGF-beta 1). This sequence and the sequence for human TGF-beta 3 (see  
 CC AAR73598) are claimed within the scope of the invention. The invention  
 CC is a composition consisting of a TGF-beta protein and an osteogenic cell  
 CC source (OCS) formulated in an acceptable carrier other than a bone  
 CC morphogenic cofactor. This composition can be used for the restoration  
 CC of bone deficiency. This provides for the generation of mature bone  
 CC only where it is required, without the inclusion of a specific  
 CC bone-inducing cofactor. This method can be used with any of the 5 human  
 CC TGF-beta's or with TGF-beta from other species.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 390 AA;

Query Match 91.0%; Score 1924; DB 16; Length 390;  
 Best Local Similarity 91.7%; Pred. No. 4.1e-162;  
 Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | : | : |
Db     61 SPSPQGEVPPGGLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | : | : | | | | | | | | | | | | | | | | | | | | | : | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | | : | | | | | | | | : | | | | |
Db    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy    361 APCCVPQALEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS 398
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    353 APCCVPQALEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 8

AAR90827

ID AAR90827 standard; Protein; 390 AA.

XX

AC AAR90827;

XX

DT 25-MAR-2003 (updated)

DT 25-JAN-1980 (first entry)

XX

DE Pre-transforming growth factor beta 1.

XX

KW transforming growth factor beta 1; wound healing;

KW recombinant production.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 8..23

FT /note= "hydrophobic domain"

FT Modified-site 82..84

FT /note= "potential N-glycosylation site"

FT Modified-site 136..138

FT /note= "potential N-glycosylation site"

FT Modified-site 176..178

FT /note= "potential N-glycosylation site"

FT Cleavage-site 277..279

FT /note= "trypsin-like peptidase cleavage site"

FT Protein 279..390

FT /label= mature\_TGF\_beta\_1

XX

PN US5482851-A.

XX

PD 09-JAN-1996.

XX

PF 05-NOV-1993; 93US-0147364.

XX

PR 13-MAR-1987; 87US-0025423.

PR 22-MAR-1985; 85US-0715142.

PR 04-AUG-1989; 89US-0389929.

PR 04-MAR-1992; 92US-0845893.

PR 05-NOV-1993; 93US-0147364.

XX

PA (GETH ) GENENTECH INC.

XX

PI Derynck RMA, Goeddel DV;

XX

DR WPI; 1996-076891/08.

DR N-PSDB; AAT15720.

XX

PT New recombinant human transforming growth factor-beta prods. - produced

PT using Chinese hamster ovary cells, for use in diagnostic applications

PT or in therapy

XX

PS Example 3; Fig 1A-C; 26pp; English.

XX

CC The pre-transforming growth factor (TGF) beta 1 protein is encoded  
 CC by AAT15720. The mature TGF beta 1 monomer is cleaved from the  
 CC precursor at the Arg-Arg dipeptide immediately preceding the mature  
 CC TGF-beta 1 NH2-terminus. It does not contain a recognisable N-terminal  
 CC signal peptide typical of most secreted proteins. The pre-TGF beta 1  
 CC contains several pairs of basic residues which could undergo  
 CC post-translational cleavage and give rise to separate polypeptide  
 CC entities. The precursor contains 3 potential N-glycosylation sites, none  
 CC of which are localised in the mature TGF beta 1. This is useful in  
 CC purification of the mature protein. TGF beta 1 can be used in, e.g. wound  
 CC healing.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 390 AA;

Query Match 91.0%; Score 1924; DB 17; Length 390;  
 Best Local Similarity 91.7%; Pred. No. 4.1e-162;  
 Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1;

Qy	1	MAPSGRLRLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAIRGQILSKRLA	60
Db	1	MPPSGRLRLPLLLPLWLLVLTTPGPPAAGLSTCKTIDMELVKKRIEAIRGQILSKRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		:   :   :	
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:   :   :	
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
		:                       :	
Db	181	YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN	300
		:	
Db	241	TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN	292
Qy	301	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	360
Db	293	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	352
Qy	361	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	398
Db	353	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 9  
 AAU77101  
 ID AAU77101 standard; Protein; 390 AA.  
 XX  
 AC AAU77101;  
 XX  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE Human transforming growth factor beta 1 (TFG-beta-1) polypeptide.

XX  
 KW Human; transforming growth factor beta; TGF-beta; insulin production;  
 KW type I diabetes mellitus; pancreatic cell outgrowth; wound healing;  
 KW pancreatic duct tissue; ischaemia; stroke; nervous system aging;  
 KW neurological condition; neurodegenerative disease; inflammation;  
 KW vasal injury; chemical injury; traumatic injury; tumour-induced injury;  
 KW amyotrophic lateral sclerosis; spinocerebellar degeneration;  
 KW immunological disease; multiple sclerosis; TGF-beta-1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200212336-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 09-FEB-2001; 2001WO-US04192.  
 XX  
 PR 09-AUG-2000; 2000US-0635368.  
 XX  
 PA (CURI-) CURIS INC.  
 XX  
 PI Wang M, Pang K;  
 XX  
 DR WPI; 2002-257468/30.  
 XX  
 PT Treating a subject with a disorder resulting from insufficient insulin  
 PT production, and inducing outgrowth of pancreatic cells, involves using  
 PT a transforming growth factor beta therapeutic -  
 XX  
 PS Disclosure; Fig 1; 77pp; English.  
 XX  
 CC The invention relates to treating a subject with a disorder resulting  
 CC from insufficient insulin production, involving contacting the subject  
 CC with a transforming growth factor beta (TGF-beta) therapeutic. TGF-beta  
 CC polypeptides can be used for treating a subject with a disorder resulting  
 CC from insufficient insulin production, e.g. type I diabetes mellitus, and  
 CC for inducing outgrowth of pancreatic cells associated with pancreatic  
 CC duct tissue within a subject. A composition comprising a TGF-beta protein  
 CC may be useful in wound healing and treatment of neurological conditions  
 CC derived from acute, subacute or chronic injury to the nervous system,  
 CC including traumatic injury, chemical injury, vasal injury and deficits  
 CC (such as ischaemia resulting from stroke), together with  
 CC infectious/inflammatory and tumour-induced injury, aging of the nervous  
 CC system including Alzheimer's disease, chronic neurodegenerative diseases  
 CC including Parkinson's disease, Huntington's chorea, amyotrophic lateral  
 CC sclerosis, spinocerebellar degenerations and chronic immunological  
 CC diseases of the nervous system or affecting the nervous system, including  
 CC multiple sclerosis. This sequence represents the human TGF-beta-1  
 CC protein.  
 XX  
 SQ Sequence 390 AA;

Query Match 90.9%; Score 1921; DB 23; Length 390;  
 Best Local Similarity 91.7%; Pred. No. 7.6e-162;  
 Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA 60

Db	1	MPPSGLRLLLLLLPLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN	300
Db	241	TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN	292
Qy	301	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	360
Db	293	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	352
Qy	361	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	398
Db	353	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 10

AAE16943

ID AAE16943 standard; Protein; 390 AA.

XX

AC AAE16943;

XX

DT 18-APR-2002 (first entry)

XX

DE Human transforming growth factor-beta1 (TGF-beta1) protein.

XX

KW Human; transforming growth factor-beta1; TGF-beta1; osteoporosis;  
 KW latency associated peptide; LAP; integrin alphavbeta3; apoptosis;  
 KW immunomodulation; inflammatory disease; fibrotic disease; cancer;  
 KW diabetic retinopathy; chronic obstructive pulmonary disorder;  
 KW bone resorption; rheumatoid arthritis; psoriasis; restenosis;  
 KW atherosclerosis; liver fibrosis; asthma; cytostatic; osteopathic;  
 KW ophthalmological; antiarteriosclerotic; vasotropic.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..29

FT /label= Signal\_peptide

FT Region 30..278

FT /note= "LAP-beta1"

FT Domain 244..246

FT /note= "RGD motif"

FT Protein 279..390

FT /note= "Human mature TGF-beta1 protein"



Db	121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAI EGFRLSAHSSSDSKDNTLHVEINGFN	240
		:                         :           :	
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWL SRGGEI EGFRLSAHCSCDSRDNTLQVDINGFT	240
Qy	241	SGRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN	300
		:	
Db	241	TGRRGDLATI HGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN	292
Qy	301	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	360
Db	293	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	352
Qy	361	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	398
Db	353	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 11

ABB82780

ID ABB82780 standard; Protein; 391 AA.

xx

AC    ABB82780;

XX

DT 18-MAR-2003 (first entry)

XX

DE TGFB1 Arg25Pro polymorphism G-allele protein sequence.

XX

KW Polymorphism; transforming growth factor beta 1; TGFB1; TGFbeta-1;

KW renal failure; nephrotropic; human; allele.

XX

OS Homo sapiens.

XX

PN WO200290585-A2.

XX

PD 14-NOV-2002.

XX

PF 08-MAY-2002; 2002WO-GB02066.

XX

PR 09-MAY-2001; 2001GB-0011277.

XX

PA (UYSH-) UNIV SHEFFIELD HALLAM.

XX

PI El-Nahas AM, Blakemore A, Khalil MS;

XX

DR WPI: 2003-120560/11.

DR N-PSDB; ABV75391.

XX

PT Determining an individual's susceptibility to the progression of renal

PT failure comprises detecting the presence of a genetic polymorphism

PT pattern in transforming growth factor beta 1 (TGFB1) gene in a sample

PT from the individual -

XX

PS Claim 51; Page 59-61; 62pp; English.

XX

CC The invention relates to determining an individual's susceptibility to

CC renal failure and involves detecting the presence of a genetic  
CC polymorphism pattern in transforming growth factor beta 1 (TGFB1) gene in  
CC a sample from an individual, where polymorphism pattern is associated  
CC with renal failure. The method is useful for determining an individual's  
CC susceptibility to the progression of renal failure. The nucleic acid  
CC comprising a T(-509)C polymorphism of TGFB1 gene, or a polypeptide  
CC comprising a sequence of 391 amino acids is useful for preparing a  
CC medicament for retarding or preventing the progression of renal disease,  
CC and for drug research purposes for retarding or preventing the  
CC progression of renal disease. Sequences ABV75386-88 represents the  
CC protein sequence for the TGFB1 G-allele of the Arg25Pro polymorphism  
CC of exon 1.

XX

SQ Sequence 391 AA;

Query Match 90.8%; Score 1920.5; DB 24; Length 391;  
Best Local Similarity 91.7%; Pred. No. 8.5e-162;  
Matches 366; Conservative 10; Mismatches 14; Indels 9; Gaps 2;

```
Qy      1 MAPSGRLRLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLRA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLRA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | : | : |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLL-RLKLKVEQHVELYQKYSNDSW 179
      | | | | | : | | : | | | | | | | | | | | | | | | | | | | | : | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRRLKLKVEQHVELYQKYSNNSW 180

Qy    180 RYLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGF 239
      | | | | | | | | | | | | | | | | | : | | | | | | | | | | | : | | | |
Db    181 RYLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGF 240

Qy    240 NSGRRGDLATIHGMNRPFLLL MATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEK 299
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TTGRRGDLATIHGMNRPFLLL MATPLERAQHLQSSRHRRALDTNYCFSST-----EK 292

Qy    300 NCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GAS 359
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    293 NCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GAS 352

Qy    360 AAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      | | | | | | | | | | | | | | | | | | | | | | | |
Db    353 AAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 391
```

RESULT 12

AAR20124

ID AAR20124 standard; Protein; 390 AA.

XX

AC AAR20124;

XX

DT 25-MAR-2003 (updated)

DT 16-APR-1992 (first entry)

XX



DE Sequence of simian transforming growth factor (TGF) beta-1.  
 XX  
 KW Hypertension therapy; hypotensive agent; blood pressure modulator.  
 XX  
 OS Monkey.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 8..21  
 FT Protein 279..390  
 XX  
 PN WO9119513-A.  
 XX  
 PD 26-DEC-1991.  
 XX  
 PF 20-JUN-1991; 91WO-US04449.  
 XX  
 PR 20-JUN-1990; 90US-0541221.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Oleson FB, Comerkeski CR;  
 XX  
 DR WPI; 1992-024199/03.  
 DR N-PSDB; AAQ20289.  
 XX  
 PT Use of transforming growth factor (TGF)-beta and their  
 PT antagonists - for modulating blood pressure, for treating  
 PT hypertension and hypotension  
 XX  
 PS Disclosure; Fig 1; 42pp; English.  
 XX  
 CC A new method for treating hypertension comprises administering a  
 CC transforming growth factor (TGF)-beta to an individual at a dose  
 CC effective for lowering blood pressure; the TGF-beta may be e.g.  
 CC mature TGF-beta, TGF-beta2, a mature TGF-beta1/beta2 hybrid, TGF-  
 CC beta1 precursor, a latent TGF-beta2 precursor, hybrid TGF-beta1/TGF-  
 CC beta2 precursor, a latent TGF-beta1 complex or a latent TGF-beta2  
 CC complex.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 390 AA;

Query Match 90.8%; Score 1919; DB 13; Length 390;  
 Best Local Similarity 91.5%; Pred. No. 1.1e-161;  
 Matches 364; Conservative 10; Mismatches 16; Indels 8; Gaps 1;

```

Qy      1 MAPSGRLRLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGRLRLPLLLPLLWLLVLTTPSRPAAGLSTCKTIDMELVKRKRIETIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180
      | | | | | : | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNNSWR 180

```

```

Qy      181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      |||:|||||:| ||||| | ||||| :|||
Db      181 YLSNRL LAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSDSKDNTLQVDINGFT 240

Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      :||||| ||||| |||
Db      241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN 292

Qy      301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      |||:|||||:| ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||
Db      293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy      361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      |||:|||||:| ||||| | ||||| | ||||| | ||||| | ||||| | |||||
Db      353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 13

AAR46227

ID AAR46227 standard; Protein; 390 AA.

XX

AC AAR46227;

XX

DT 25-MAR-2003 (updated)

DT 09-JUL-1994 (first entry)

XX

DE Human pre-TGF-beta-1.

XX

KW TGF-beta-1; TGF-beta-2; transforming growth factor beta-1;

KW transforming growth factor beta-3; recombinant; wound healing;

KW vulnerable.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 279..390

FT /label= Mat\_peptide

FT Cleavage-site 279

FT /note= "TGF-beta-1 release site"

FT Modified-site 82..84

FT /label= N-glycosylation\_site

FT Modified-site 136..138

FT /label= N-glycosylation\_site

FT Modified-site 176..178

FT /label= N-glycosylation\_site

XX

PN US5284763-A.

XX

PD 08-FEB-1994.

XX

PF 04-MAR-1992; 92US-0845893.

XX

PR 22-MAR-1985; 85US-0715142.

PR 13-MAR-1987; 87US-0025423.

PR 04-AUG-1989; 89US-0389929.

PR 04-MAR-1992; 92US-0845893.

```
XX      (GETH ) GENENTECH INC.
XX
PI      Derynk RMA,   Goeddel DV;
XX
DR      WPI; 1994-056343/07.
DR      N-PSDB; AAQ56923.
XX
PT      Nucleic acid sequences encoding transforming growth factor-beta -
PT      diagnostic probes, and for use in therapeutics
XX
PS      Disclosure; Fig 1b; 25pp; English.
XX
CC      cDNA sequences were determined for human pre-TGF-beta-1 (AAQ56923),
CC      pig TGF-beta-3 (AAQ56925) and human TGF-beta-3 (AAQ56926), and the
CC      corresponding amino acid sequences were determined (AAR46227-29,
CC      respectively). A genomic fragment corresponding to a human TGF-
CC      beta-1 exon (AAQ56924) was also isolated and its amino acid sequence
CC      determined (AAR46230). The sequences have been used in the
CC      construction of vectors for the expression of recombinant TGF-
CC      beta.
CC      (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ      Sequence      390 AA;

Query Match          90.6%; Score 1915; DB 15; Length 390;
Best Local Similarity 91.5%; Pred. No. 2.6e-161;
Matches 364; Conservative 10; Mismatches 16; Indels 8; Gaps 1;

Qy           1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLRA 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db           1 MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLRA 60

Qy           61 SPSPSGQDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
| | | | |: | | | | | | | | | | | | | | | | | | | | | | | | | | | : | : |
Db           61 SPSPSGQEVPPGPLPEAVLALYNSTRDRVAGESAEREPEPEADYYAKEVTRVLMVETHNEI 120

Qy           121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKEQHVELYQKYSNDSWR 180
| | | | : | | : | | | | | | | | | | | | | | | | | | | | | | | | : | | |
Db           121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLVKEQHVELYQKYSNNSWR 180

Qy           181 YLSNRLLAPSDSPEWL SFDVTGVVRQWL TRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | : | | | |
Db           181 YLSNRLLAPSDSPEWL SFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240

Qy           241 SGRRGDLATIHGMNRPFLLL MATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |||
Db           241 TGRRGDLATIHGMNRPFLLL MATPLERAQHLQSSRHRRALDTNYCFSST-----EKN 292

Qy           301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKV LALYNQHNP GASA 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db           293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKV LALYNQHNP GASA 352

Qy           361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db           353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 14

AAR83054

ID AAR83054 standard; Protein; 391 AA.

XX

AC AAR83054;

XX

DT 25-JUN-1996 (first entry)

XX

DE Transforming growth factor-beta 1.

XX

KW macrophage inducible nitric oxide synthase; iNOS; constitutive NOS;

KW interleukin-1-beta; transforming growth factor-beta; TGF-beta; IL1-beta;

KW nitric oxide production; hypotension; inflammation; septic shock;

KW treatment.

XX

OS Mammalian sp.

XX

FH Key Location/Qualifiers

FT Protein 279..391

FT /note= "represents the mature active TGF beta-1 mol."

XX

PN W09526745-A1.

XX

PD 12-OCT-1995.

XX

PF 05-APR-1994; 94WO-US03705.

XX

PR 05-APR-1994; 94WO-US03705.

XX

PA (HARD ) HARVARD COLLEGE.

XX

PI Lee M, Perrella MA;

XX

DR WPI; 1995-358443/46.

DR N-PSDB; AAT05876.

XX

PT Treatment of hypotension, esp. in septic shock - by administering

PT transforming growth factor-beta e.g. to inhibit inducible nitric

PT oxide synthase gene transcription

XX

PS Disclosure; Fig 17; 52pp; English.

XX

CC Transforming growth factor-beta 1 (TGF-beta 1) has been found to inhibit

CC inducible nitric oxide synthase (iNOS) gene transcription, esp. in

CC interleukin-1-beta (IL1-beta) stimulated rat smooth muscle cells, and at

CC a dose which does not inhibit constitutive NOS. TGF-beta 1 or 2

CC (AAR83055) or their active fragments (esp. derived from the

CC carboxy-terminal 112 amino acids), can be used in the treatment of

CC hypotension, such as that associated with severe inflammation or septic

CC shock.

XX

SQ Sequence 391 AA;

Query Match 90.5%; Score 1913.5; DB 16; Length 391;

Best Local Similarity 91.5%; Pred. No. 3.5e-161;

Matches 365; Conservative 10; Mismatches 15; Indels 9; Gaps 2;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA 60

Qy     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | : | |
Db     61 SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLL-RLKLKVEQHVELYQKYSNDSW 179
      | | | | : | | : | | | | | | | | | | | | | | | | | | | : | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRRLKLKVEQHVELYQKYSNNSW 180

Qy    180 RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGF 239
      | | | | | | | | | | | | | | | | : | | | | | | | | : | | | |
Db    181 RYLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGF 240

Qy    240 NSGRRGDLATIHGMRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEK 299
      : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TTGRRGDLATIHGMRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EK 292

Qy    300 NCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPAS 359
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    293 NCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPAS 352

Qy    360 AAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      | | | | | | | | | | | | | | | | | | | | | | | | |
Db    353 AAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 391

```

RESULT 15

AAW78785

ID AAW78785 standard; Protein; 390 AA.

XX

AC AAW78785;

XX

DT 25-MAR-2003 (updated)

DT 21-DEC-1998 (first entry)

XX

DE Human pre-transforming growth factor-beta 1.

XX

KW Transforming growth factor-beta 1; TGF-beta 1; human.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 8..23

FT /note= "hydrophobic domain"

FT Protein 279..390

FT /label= Mat\_protein

FT Modified-site 82..84

FT /note= "Asn is N-glycosylated"

FT Modified-site 136..138

FT /note= "Asn is N-glycosylated"

FT Modified-site 176..178

FT /note= "Asn is N-glycosylated"

FT Cleavage-site 277..278

```

FT      /note= "cleavage site for release of TGF-beta 1"
XX
PN      US5801231-A.
XX
PD      01-SEP-1998.
XX
PF      30-MAY-1995;      95US-0454468.
XX
PR      13-MAR-1987;      87US-0025423.
PR      22-MAR-1985;      85US-0715142.
PR      04-AUG-1989;      89US-0389929.
PR      04-MAR-1992;      92US-0845893.
PR      05-NOV-1993;      93US-0147364.
PR      30-MAY-1995;      95US-0454468.
XX
PA      (GETH ) GENENTECH INC.
XX
PI      Derynck RMA,  Goeddel DV;
XX
DR      WPI; 1998-494840/42.
DR      N-PSDB; AAV52933.
XX
PT      DNA encoding transforming growth factor-beta precursor sequence -
PT      useful for analysis to perform manipulations to increase yield of
PT      recombinant production of the protein
XX
PS      Example 3; Fig 1B 1-3; 26pp; English.
XX
CC      This is the amino acid sequence of human transforming growth
CC      factor-beta 1 precursor (preTGF-beta 1). It was deduced from
CC      a preTGF-beta 1 cDNA sequence (see AAV52933). The invention relates
CC      to the recombinant production of TGF-beta. Biologically active
CC      TGF-beta is defined as being capable of inducing EGF-potentiated
CC      anchorage independent growth of target cell lines and/or growth
CC      inhibition of neoplastic cell lines. Nucleic acids encoding
CC      TGF-beta have been isolated and cloned into vectors which are
CC      replicated in bacteria and expressed in eukaryotic cells. TGF-beta
CC      recovered from transformed cells is used in known therapeutic
CC      applications.
CC      (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ      Sequence      390 AA;

Query Match          90.5%;  Score 1913;  DB 19;  Length 390;
Best Local Similarity 91.2%;  Pred. No. 3.9e-161;
Matches 363;  Conservative 10;  Mismatches 17;  Indels 8;  Gaps 1;

QY      1  MAPSGLRLLPLLLPLLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1  MPPSGLRLLPLLLPLLWLLVLTPGPPAPGLSTCKTIDMEQVKRKRIEAIRGQILSKLRLA 60

QY      61  SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
        | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61  SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

QY      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKEQHVELYQKYSNDSWR 180
        | | | | | : | | : | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db      121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180
Qy      181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
        |||
Db      181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240
        |||
Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
        :|||
Db      241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN 292
        |||
Qy      301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
        |||
Db      293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352
        |||
Qy      361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
        |||
Db      353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

Search completed: October 28, 2003, 09:06:47  
Job time : 34.2026 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 09:09:54 ; Search time 24.1473 Seconds  
(without alignments)  
2760.110 Million cell updates/sec

Title: US-10-017-372E-13  
Perfect score: 2114  
Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 398

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*

```

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	2035	96.3	390	11	US-09-214-592-26	Sequence 26, Appl
2	1934	91.5	390	11	US-09-214-592-29	Sequence 29, Appl
3	1931	91.3	390	15	US-10-087-268-2	Sequence 2, Appli
4	1924	91.0	390	15	US-10-087-268-5	Sequence 5, Appli
5	1921	90.9	390	12	US-10-276-947-1	Sequence 1, Appli
6	1919	90.8	390	11	US-09-214-592-33	Sequence 33, Appl
7	1913.5	90.5	391	11	US-09-214-592-17	Sequence 17, Appl
8	1910	90.4	390	10	US-09-756-283A-23	Sequence 23, Appl
9	1898	89.8	390	11	US-09-214-592-28	Sequence 28, Appl
10	1840	87.0	390	11	US-09-214-592-20	Sequence 20, Appl
11	1840	87.0	390	11	US-09-214-592-23	Sequence 23, Appl
12	1577	74.6	315	11	US-09-214-592-25	Sequence 25, Appl
13	1262.5	59.7	455	10	US-09-756-283A-20	Sequence 20, Appl
14	1139	53.9	447	10	US-09-756-283A-22	Sequence 22, Appl
15	1063.5	50.3	373	11	US-09-214-592-32	Sequence 32, Appl
16	876	41.4	412	11	US-09-214-592-31	Sequence 31, Appl
17	874	41.3	382	11	US-09-214-592-34	Sequence 34, Appl
18	873.5	41.3	383	10	US-09-756-283A-27	Sequence 27, Appl
19	838.5	39.7	409	11	US-09-214-592-27	Sequence 27, Appl
20	832	39.4	410	11	US-09-214-592-22	Sequence 22, Appl
21	828	39.2	412	11	US-09-214-592-24	Sequence 24, Appl
22	826	39.1	412	11	US-09-214-592-19	Sequence 19, Appl
23	826	39.1	412	14	US-10-028-158-21	Sequence 21, Appl
24	822	38.9	412	10	US-09-756-283A-25	Sequence 25, Appl
25	821.5	38.9	414	11	US-09-214-592-21	Sequence 21, Appl
26	808.5	38.2	304	10	US-09-756-283A-26	Sequence 26, Appl
27	806.5	38.2	414	10	US-09-756-283A-24	Sequence 24, Appl
28	806.5	38.2	414	11	US-09-214-592-18	Sequence 18, Appl
29	805.5	38.1	412	11	US-09-214-592-30	Sequence 30, Appl
30	758	35.9	139	14	US-10-002-278-8	Sequence 8, Appli
31	642	30.4	115	10	US-09-859-211-47	Sequence 47, Appl
32	642	30.4	115	10	US-09-880-708-25	Sequence 25, Appl
33	642	30.4	115	11	US-09-872-856-47	Sequence 47, Appl
34	642	30.4	115	15	US-10-335-483-29	Sequence 29, Appl
35	634	30.0	114	10	US-09-813-459-22	Sequence 22, Appl



36	634	30.0	114	14	US-10-115-406-21	Sequence 21, Appl
37	634	30.0	114	15	US-10-154-333-23	Sequence 23, Appl
38	624	29.5	112	10	US-09-813-271B-2	Sequence 2, Appli
39	624	29.5	113	10	US-09-813-398-13	Sequence 13, Appl
40	562	26.6	98	12	US-10-187-394-1	Sequence 1, Appli
41	548	25.9	116	14	US-10-115-406-24	Sequence 24, Appl
42	548	25.9	116	15	US-10-154-333-26	Sequence 26, Appl
43	521	24.6	112	10	US-09-813-271B-8	Sequence 8, Appli
44	499	23.6	98	12	US-10-187-394-4	Sequence 4, Appli
45	497	23.5	114	14	US-10-115-406-25	Sequence 25, Appl

## ALIGNMENTS

## RESULT 1

```

US-09-214-592-26
; Sequence 26, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
;   APPLICANT: Yamasaki,CMotoo
;   APPLICANT: Shibata,CKenji
;   APPLICANT: Sato,CYasufumi
;   TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
;   TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
;   FILE REFERENCE: 11060
;   CURRENT APPLICATION NUMBER: US/09/214,592A
;   CURRENT FILING DATE: 1999-01-18
;   NUMBER OF SEQ ID NOS: 34
;   SOFTWARE:
; SEQ ID NO 26
;   LENGTH: 390
;   TYPE: PRT
;   ORGANISM: porcine
US-09-214-592-26

```

Query Match 96.3%; Score 2035; DB 11; Length 390;  
Best Local Similarity 97.0%; Pred. No. 7.2e-180;  
Matches 386; Conservative 1; Mismatches 3; Indels 8; Gaps 1;

[illegible]

Qy	241	SGRRGDLATIHGMRNPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN	300
Db	241	SGRRGDLATIHGMRNPFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN	292
Qy	301	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPGLASA	360
Db	293	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPGLASA	352
Qy	361	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	398
Db	353	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

## RESULT 2

US-09-214-592-29

; Sequence 29, Application US/09214592A

; Publication No. US20030027218A1

; GENERAL INFORMATION:

; APPLICANT: Yamasaki, CMotoo

; APPLICANT: Shibata,CKenji

; APPLICANT: Sato, CYasufumi

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND

## METHOD

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

; FILE REFERENCE: 11060

; CURRENT APPLICATION NUMBER: US/09/214,592A

; CURRENT FILING DATE: 1999-01-18

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE:

; SEQ ID NO 29

; LENGTH: 390

; TYPE: PRT

; ORGANISM: ovine

US-09-214-592-29

Query Match 91.5%; Score 1934; DB 11; Length 390;

Best Local Similarity 91.7%; Pred. No. 1.6e-170;

Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1;

[illegible]

```

Db      241 SGRRGDLATIHGMRNPFLLLMATPLERAQHLHSSRRRALDTNYCFSST-----EKN 292
QY      301 CCVRQLYIDFRKDLGWKIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
        |||||||
Db      293 CCVRQLYIDFRKDLGWKIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352
QY      361 APCCVPQALEPLPIVYYYVGRKPKEQLSNMIVRSCKCS 398
        |||||||
Db      353 APCCVPQALEPLPIVYYYVGRKPKEQLSNMIVRSCKCS 390

```

Query Match 91.3%; Score 1931; DB 15; Length 390;  
Best Local Similarity 92.0%; Pred. No. 3.1e-170;  
Matches 366; Conservative 10; Mismatches 14; Indels 8; Gaps 1;

```

Db      293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352
Qy      361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
        |||||||||||||||||||||||||||||||||||||||
Db      353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 4

US-10-087-268-5

```

; Sequence 5, Application US/10087268
; Publication No. US20030119010A1
; GENERAL INFORMATION:
; APPLICANT: Jonsonn, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or
condition
; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Human
US-10-087-268-5

```

```

Query Match          91.0%; Score 1924; DB 15; Length 390;
Best Local Similarity 91.7%; Pred. No. 1.4e-169;
Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1;

```

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKKRIEAI RQILSKRLA 60
        | ||||||||||||||||||| |||||||||||||||||||||||
Db      1 MPPSGLRLLPLLLPLLWLLVLTGP PPAAGLSTCKTIDMELVKKRIEAI RQILSKRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
        |||||: ||:| ||||||||||||||||||| |||||||||||||||: ||
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
        |||||||||||||||||||:| ||||||| | ||: ||| |: |||
Db    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSDSRDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
        : ||||||||||||||||||| ||||||||||||||| ||
Db    241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
        |||||||||||||||||||||||||||||||||||||||
Db    293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy    361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
        |||||||||||||||||||||||||||||||||||

```



RESULT 6

US-09-214-592-33

; Sequence 33, Application US/09214592A

; Publication No. US20030027218A1

; GENERAL INFORMATION:

; APPLICANT: Yamasaki,CMotoo

; APPLICANT: Shibata,CKenji

; APPLICANT: Sato,CYasufumi

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND  
METHOD

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

; FILE REFERENCE: 11060

; CURRENT APPLICATION NUMBER: US/09/214,592A

; CURRENT FILING DATE: 1999-01-18

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE:

; SEQ ID NO 33

; LENGTH: 390

; TYPE: PRT

; ORGANISM: simian

US-09-214-592-33

Query Match 90.8%; Score 1919; DB 11; Length 390;

Best Local Similarity 91.5%; Pred. No. 4e-169;

Matches 364; Conservative 10; Mismatches 16; Indels 8; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKKRIEAI RQILSKRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKKRIETIRGQILSKRLA 60

Qy     61 SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | : | : |
Db     61 SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180
      | | | | : | | : | | | | | | | | | | | | | | | | | | | | | : | | |
Db    121 YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | : | | | | | | | | | | | | | | | | | | | | | : | | |
Db    181 YLSNRL LAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSDSKDNTLQVDINGFT 240

Qy    241 SGRRGDLATI HGMNRPFLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATI HGMNRPFLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy    361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 7

US-09-214-592-17  
; Sequence 17, Application US/09214592A  
; Publication No. US20030027218A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamasaki,CMotoo  
; APPLICANT: Shibata,CKenji  
; APPLICANT: Sato,CYasufumi  
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND  
METHOD  
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS  
; FILE REFERENCE: 11060  
; CURRENT APPLICATION NUMBER: US/09/214,592A  
; CURRENT FILING DATE: 1999-01-18  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE:  
; SEQ ID NO 17  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: human  
US-09-214-592-17

Query Match 90.5%; Score 1913.5; DB 11; Length 391;  
Best Local Similarity 91.5%; Pred. No. 1.3e-168;  
Matches 365; Conservative 10; Mismatches 15; Indels 9; Gaps 2;

Qy	1	MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAI	RQGILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLTTPGPPAAGLSTCKTIDMELVKKRIEAI	RQGILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEV	TRVLMVESGNQI	120
		:		
Db	61	SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEV	TRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLL-RLKLKVEQH	VELYQKYSNDSW	179
		:     :		
Db	121	YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRRLKLKVEQH	VELYQKYSNNNSW	180
Qy	180	RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDS	KDNTLHVEINGF	239
		:		
Db	181	RYLSNRLLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCD	SRDNTLQVDINGF	240
Qy	240	NSGRRGDLATI HGMNRPFLLL MATPLERAQHLHSSRHRRALDTNYC	FSSTDYKDDDDKEK	299
		:		
Db	241	TTGRRGDLATI HGMNRPFLLL MATPLERAQHLQSSRHRRALDTNYC	FSST-----EK	292
Qy	300	NCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK	VLALYNQHNP GAS	359
Db	293	NCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK	VLALYNQHNP GAS	352
Qy	360	AAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS		398
Db	353	AAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS		391

RESULT 8  
US-09-756-283A-23  
; Sequence 23, Application US/09756283A

```
; Patent No. US20020151478A1
; GENERAL INFORMATION:
; APPLICANT: Chernajovsky, Yuti
; APPLICANT: Dreja, Hanna Stina
; APPLICANT: Adams, Gillian
; TITLE OF INVENTION: Latent Fusion Protein
; FILE REFERENCE: 0623.1000000
; CURRENT APPLICATION NUMBER: US/09/756,283A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-756-283A-23
```

```
Query Match          90.4%; Score 1910; DB 10; Length 390;
Best Local Similarity 91.2%; Pred. No. 2.7e-168;
Matches 363; Conservative 10; Mismatches 17; Indels 8; Gaps 1;
```

```
Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | : : |
Db     61 SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHHEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | : | |
Db    121 YDKFKQSTHSTYMFNISELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | : | | | | | | | | | | | | | | | | | |
Db    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLL MATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMNRPFLLL MATPLERAQHLQSSRHRRALDTNYCFSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy    361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

```
RESULT 9
US-09-214-592-28
; Sequence 28, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
; APPLICANT: Yamasaki, CMotoo
; APPLICANT: Shibata, CKenji
```





```
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 20
; LENGTH: 390
; TYPE: PRT
; ORGANISM: murine
US-09-214-592-20
```

```
Query Match          87.0%; Score 1840; DB 11; Length 390;
Best Local Similarity 86.9%; Pred. No. 8.3e-162;
Matches 346; Conservative 15; Mismatches 29; Indels 8; Gaps 1;
```

```
Qy      1 MAPSGLRLLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA 60
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 MPPSGLRLLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA 60

Qy     61 SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     61 SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      |:| | ||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    121 YEKTKDISHSIYMFNTSDIREAVPEPPLLSRAELRLQRLKSSVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      || |||| |:|:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    181 YLGNRLLTPTDTPWLSFDVTGVVRQWLNQGDGIQGFRFSAHCSCDSKDNKLHVEINGIS 240

Qy    241 SGRRGDLATIHGMRNPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    241 PKRRGDLGTIHDNRNPFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy    361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    353 SPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

# RESULT 11

```
US-09-214-592-23
; Sequence 23, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
; APPLICANT: Yamasaki,CMotoo
; APPLICANT: Shibata,CKenji
; APPLICANT: Sato,CYasufumi
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
```



; CURRENT APPLICATION NUMBER: US/09/214,592A  
; CURRENT FILING DATE: 1999-01-18  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE:  
; SEQ ID NO 25  
; LENGTH: 315  
; TYPE: PRT  
; ORGANISM: bovine  
US-09-214-592-25

Query Match 74.6%; Score 1577; DB 11; Length 315;  
Best Local Similarity 91.3%; Pred. No. 1.4e-137;  
Matches 295; Conservative 9; Mismatches 11; Indels 8; Gaps 1;

```
Qy      76 AVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLF 135
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1  AILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKIYDKMKSSSHSIYMF 60

Qy     136 NTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWRYLSNRL LAPSDSPEW 195
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 NTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWRYLSNRL LAPSDSPEW 120

Qy     196 LSFDTVGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFNSGRRGDLATIHGMNR 255
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     121 LSFDTVGVVRQWLTRREEIEGFRLSAHCSCDSKDNTLQVDINGFSSGRRGDLATIHGMNR 180

Qy     256 PFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLG 315
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     181 PFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKNCCVRQLYIDFRKDLG 232

Qy     316 WKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIV 375
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     233 WKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIV 292

Qy     376 YYVGRKPKVEQLSNMIVRSCKCS 398
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     293 YYVGRKPKVEQLSNMIVRSCKCS 315
```

RESULT 13

US-09-756-283A-20  
; Sequence 20, Application US/09756283A  
; Patent No. US20020151478A1  
; GENERAL INFORMATION:  
; APPLICANT: Chernajovsky, Yuti  
; APPLICANT: Dreja, Hanna Stina  
; APPLICANT: Adams, Gillian  
; TITLE OF INVENTION: Latent Fusion Protein  
; FILE REFERENCE: 0623.1000000  
; CURRENT APPLICATION NUMBER: US/09/756,283A  
; CURRENT FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 20  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Artificial Sequence



Db	190	PLGLW---AGGGSAALSTCKTIDMELVKRKRIEAIRGQILSKLRLASPPSQGEVPPGP	245
Qy	73	LPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMMVESGNQIYDKFKGTPHSLY	132
		:   :         :     :	
Db	246	LPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMMVETHNEIYDKFKQSTHSIY	305
Qy	133	MLFNTSELREAVPEPVLLSRAELRLL-RLKLKVEQHVELYQKYNSDSWRYLSNRLLAPSD	191
		:	
Db	306	MFFNTSELREAVPEPVLLSRAELRLLRRLLKLKVEQHVELYQKYNNDSWRYLSNRLLAPSD	365
Qy	192	SPEWL SFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN SGRRGDLATIH	251
		:                   :     :       :	
Db	366	SPEWL SFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFTTGRRGDLATIH	425
Qy	252	GMNRPFLLLMATPLERAQHLHS	273
Db	426	GMNRPFLLLMATPLERAQHLQS	447

US-09-214-592-32

; Sequence 32, Application US/09214592A

: GENERAL INFORMATION:

; APPLICANT: Shibata,CKenji

10 TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

CURRENT APPLICATION NUMBER: US/09/214,592A

NUMBER OF SEO ID NOS: 34

; SEO ID NO 32

; TYPE: PRT

US-09-214-592-32

Best Local Similarity 54.9%; Pred. No. 6.2e-90;

Matches 209; Conservative 51; Mismatches 100; Indels 21; Gaps 8;

QY 30 LSTCKTIDMELVKRKRIEAI RQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVA 89  
| | | | : : : | : | | | : | | : | | : | | | | : : :

QY           90 GES-VEPEPEPEADYYAKEVTRVL MVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPV 148  
               : : | : : : ||| : : | : : : : : : : ||| : |

QY 149 LLSRAELRLRLRLK-----VEQHVELYQKYSNDSWRYLSNRL LAPSDSPEWLSFDVTGV 203

```

Qy      204 VRQWLTRREAIEGFRLSAHSSSD--SKDNTLHVEINGFNSGRRGDLATIHGMNR--PFL 258
      | |||: | : ||| | : : || :|||: | : | ||:
Db      181 VHQWLSGSELLGVFKLSVHCPCEMGPGHAEEMRISIEGFEQ-QRGDMQSI AKKHRRVPYV 239

Qy      259 LLMATPLERAQHLHSSRRRALDTNYCFS-STDYKDDDDKEKNCCVRQLYIDFRKDLGWK 317
      | || | ||| |||:| || |||:||| || | ||||| ||||| |||
Db      240 LAMALPAERANELHSARRRRDLDTDYCFGPGTD-----EKNCCVRPLYIDFRKDLQWK 292

Qy      318 WIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYY 377
      ||||| |||: ||||| ||||| ||||| ||||| ||||| ||||| |||
Db      293 WIHEPKGYMANFCMGPCPYIWSADTQYTKVLALYNQHNP GASAAPCCVPQTL DPLPIIYY 352

Qy      378 VGRKPKVEQLSNMIVRSCKCS 398
      ||| : |||||: ||: |||
Db      353 VGRNVRVEQLSNMVVRACKCS 373

```

Search completed: October 28, 2003, 09:28:00  
Job time : 25.1473 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 00:37:16 ; Search time 13.1517 Seconds  
(without alignments)  
2910.285 Million cell updates/sec

Title: US-10-017-372E-13  
Perfect score: 2114  
Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 398

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	2038	96.4	390	2	A27512	transforming growt	
2	2006.5	94.9	391	2	S01413	transforming growt	
3	1934	91.5	390	2	I46463	transforming growt	
4	1921	90.9	390	1	WFHU2	transforming growt	
5	1919	90.8	390	2	A26960	transforming growt	
6	1898	89.8	390	2	JC4023	transforming growt	
7	1840	87.0	390	1	WFMS2	transforming growt	
8	1840	87.0	390	2	S10219	transforming growt	
9	1577	74.6	315	2	A40057	transforming growt	
10	1063.5	50.3	373	2	A41918	transforming growt	
11	876	41.4	412	2	A34939	transforming growt	
12	874	41.3	382	2	B61036	transforming growt	
13	838.5	39.7	409	2	S01825	transforming growt	
14	832	39.4	410	2	A41397	transforming growt	
15	828	39.2	410	2	A55706	transforming growt	
16	826	39.1	412	2	A36169	transforming growt	
17	821.5	38.9	414	1	WFMSB2	transforming growt	
18	806.5	38.2	414	1	WFMKB2	transforming growt	
19	806.5	38.2	414	2	A31249	transforming growt	
20	806	38.1	413	1	WFXLB2	transforming growt	
21	805.5	38.1	412	2	A39489	transforming growt	
22	790.5	37.4	442	2	B31249	transforming growt	
23	696	32.9	130	2	I48196	transforming growt	
24	468	22.1	112	2	A61439	transforming growt	
25	289.5	13.7	425	2	I47072	inhibin beta-A cha	
26	287	13.6	424	1	WFPGBA	inhibin beta-A cha	
27	285	13.5	424	1	S31440	inhibin beta-A cha	
28	284.5	13.5	425	1	S50898	inhibin beta-A cha	
29	283	13.4	426	1	B24248	inhibin beta-A cha	
30	281	13.3	424	1	B40905	inhibin beta-A cha	
31	268	12.7	413	2	JC4862	activin beta-A cha	
32	264	12.5	398	2	JH0688	bone morphogenetic	
33	262.5	12.4	398	2	JH0687	bone morphogenetic	
34	257.5	12.2	394	2	S45355	bone morphogenetic	
35	256.5	12.1	396	1	BMHU2	bone morphogenetic	
36	255	12.1	513	1	BMHU6	bone morphogenetic	
37	252.5	11.9	510	2	A54798	Vg-1-related prote	
38	249	11.8	393	2	S37073	bone morphogenetic	
39	243.5	11.5	350	2	JC5241	activin beta E cha	
40	241.5	11.4	420	2	I49541	bone morphogenetic	
41	238.5	11.3	455	2	A43918	TGF-beta-related p	
42	236	11.2	461	2	S52408	SPDVR1 protein - s	
43	234.5	11.1	408	2	JH0801	bone morphogenetic	
44	233.5	11.0	367	2	JC4151	activin beta D cha	
45	233.5	11.0	407	1	A40150	inhibin beta-B cha	

## ALIGNMENTS

RESULT 1



A27512  
transforming growth factor beta-1 precursor - pig  
N;Alternate names: TGF-beta  
C;Species: *Sus scrofa domestica* (domestic pig)  
C;Date: 05-Jun-1988 #sequence\_revision 05-Jun-1988 #text\_change 16-Jul-1999  
C;Accession: A27512; A26356; I46657  
R;Derynck, R.; Rhee, L.  
Nucleic Acids Res. 15, 3187, 1987  
A;Title: Sequence of the porcine transforming growth factor-beta precursor.  
A;Reference number: A27512; MUID:87174844; PMID:3470708  
A;Accession: A27512  
A;Molecule type: mRNA  
A;Residues: 1-390 <DER>  
R;Cheifetz, S.; Weatherbee, J.A.; Tsang, M.L.S.; Anderson, J.K.; Mole, J.E.;  
Lucas, R.; Massague, J.  
Cell 48, 409-415, 1987  
A;Title: The transforming growth factor-beta system, a complex pattern of cross-  
reactive ligands and receptors.  
A;Reference number: A90890; MUID:87102890; PMID:2879635  
A;Accession: A26356  
A;Molecule type: protein  
A;Residues: 279-322 <CHE>  
R;Kondaiah, P.; Van Obberghen-Schilling, E.; Ludwig, R.L.; Dhar, R.; Sporn,  
M.B.; Roberts, A.B.  
J. Biol. Chem. 263, 18313-18317, 1988  
A;Title: cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.  
Evidence for alternate splicing and polyadenylation.  
A;Reference number: I46657; MUID:89054010; PMID:2461367  
A;Accession: I46657  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-390 <KON>  
A;Cross-references: GB:M23703; NID:g755044; PIDN:AAA64616.1; PID:g755045  
C;Genetics:  
A;Gene: TGFB; TGF-beta-1  
C;Superfamily: inhibin  
C;Keywords: growth factor

Query Match 96.4%; Score 2038; DB 2; Length 390;  
Best Local Similarity 97.2%; Pred. No. 8e-155;  
Matches 387; Conservative 0; Mismatches 3; Indels 8; Gaps 1;

Qy	1	MAPSGRLRLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLA	60
Db	1	MPPSGRLRLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLA	60
Qy	61	SPPSQGDVPFGPGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMEVSGNQI	120
Db	61	SPPSQGDVPFGPGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMEVSGNQI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Db	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
Qy	181	YLSNRLLAPSDSPEWLSEFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
Db	181	YLSNRLLAPSDSPEWLSEFDVTGVVRQWLTRREAIEGFRLSAHCSCDSKDNTLHVEINGFN	240

```

Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      |||
Db      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN 292

Qy      301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      |||
Db      293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy      361 APCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 398
      |||
Db      353 APCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

```

## RESULT 2

S01413

transforming growth factor beta-1 precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 24-Nov-1999

C;Accession: S01413

R;Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B.

Nucleic Acids Res. 16, 8730, 1988

A;Title: Nucleotide sequence of chicken transforming growth factor-beta 1 (TGF-beta 1).

A;Reference number: S01413; MUID:88335639; PMID:3166520

A;Accession: S01413

A;Molecule type: DNA

A;Residues: 1-391 <JAK>

A;Cross-references: EMBL:X12373; NID:g63808; PIDN:CAA30933.1; PID:g63809

C;Superfamily: inhibin

C;Keywords: growth factor

```

Query Match          94.9%;  Score 2006.5;  DB 2;  Length 391;
Best Local Similarity 96.2%;  Pred. No. 2.6e-152;
Matches 384;  Conservative 0;  Mismatches 6;  Indels 9;  Gaps 2;

```

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RQILSKRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGPGLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RQILSKRLA 60

Qy      61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120

Qy      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWG 180

Qy      181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEIN-GF 239
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHCSDSKDNTLHVEINAGF 240

Qy      240 NSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEK 299
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      241 NSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EK 292

Qy      300 NCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GAS 359

```



RESULT 4

WFHU2

transforming growth factor beta-1 precursor [validated] - human

N;Alternate names: growth-inhibitory factor; TGF type 2; TGF-beta

C;Species: Homo sapiens (man)

C;Date: 28-Feb-1986 #sequence\_revision 19-Oct-1995 #text\_change 08-Dec-2000

C;Accession: A27513; A01395; A22290; I59664; S53444

R;Derynck, R.; Rhee, L.; Chen, E.Y.; Van Tilburg, A.

Nucleic Acids Res. 15, 3188-3189, 1987

A;Title: Intron-exon structure of the human transforming growth factor-beta precursor gene.

A;Reference number: A27513; MUID:87174845; PMID:3470709

A;Accession: A27513

A;Molecule type: DNA

A;Residues: 1-390 <DER>

A;Cross-references: GB:X05839; GB:Y00112; NID:g37097; PIDN:CAA29283.1;

PID:g1212989

R;Derynck, R.; Jarrett, J.A.; Chen, E.Y.; Eaton, D.H.; Bell, J.R.; Assoian, R.K.; Roberts, A.B.; Sporn, M.B.; Goeddel, D.V.

Nature 316, 701-705, 1985

A;Title: Human transforming growth factor-beta complementary DNA sequence and expression in normal and transformed cells.

A;Reference number: A01395; MUID:85296301; PMID:3861940

A;Accession: A01395

A;Molecule type: mRNA

A;Residues: 1-9, 'P',11-24, 'P',26-159, 'R',160-390 <DE2>

A;Cross-references: GB:X02812; GB:J05114; NID:g37092; PIDN:CAA26580.1;

PID:g37093

A;Note: the authors suggest that residues 8-23 could represent the hydrophobic core of an amino-terminal signal peptide

R;Massague, J.; Like, B.

J. Biol. Chem. 260, 2636-2645, 1985

A;Title: Cellular receptors for type beta transforming growth factor. Ligand binding and affinity labeling in human and rodent cell lines.

A;Reference number: A22290; MUID:85131019; PMID:2982829

A;Accession: A22290

A;Molecule type: protein

A;Residues: 279-295, 'XX',298-301 <MAS>

R;Urushizaki, Y.; Niitsu, Y.; Terui, T.; Koshida, Y.; Mahara, K.; Kohgo, Y.; Urushizaki, I.; Takahashi, Y.; Ito, H.

Tumor Res. 22, 41-55, 1987

A;Title: Cloning and expression of the gene for human transforming growth factor-beta in Escherichia coli.

A;Reference number: I59664

A;Accession: I59664

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 279-390 <RES>

A;Cross-references: GB:M38449; NID:g339557; PIDN:AAA36735.1; PID:g339558

R;Stam, K.; Stewart, A.A.; Qu, G.Y.; Iwata, K.K.; Fenyo, D.; Chait, B.T.; Marshak, D.R.; Haley, J.D.

Biochem. J. 305, 87-92, 1995

A;Title: Physical and biological characterization of a growth-inhibitory activity purified from the neuroepithelioma cell line A673.

A;Reference number: S53444; MUID:95126934; PMID:7826358  
A;Accession: S53444  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 279-297 <STA>  
C;Comment: The mature protein is the carboxyl-terminal segment of a precursor polypeptide; the active molecule is a dimer of identical polypeptide chains linked by an interchain disulfide bond.  
C;Genetics:  
A;Gene: GDB:TGFB1; TGFB  
A;Cross-references: GDB:120729; OMIM:190180  
A;Map position: 19q13.2-19q13.2  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; homodimer; mitogen; transformation  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-278/Domain: propeptide #status predicted <PRO>  
F;244-246/Region: cell attachment (R-G-D) motif  
F;279-390/Product: transforming growth factor beta-1 #status experimental <MAT>  
F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.9%; Score 1921; DB 1; Length 390;  
Best Local Similarity 91.7%; Pred. No. 1.7e-145;  
Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1;

```
Qy      1 MAPSGLRLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLLLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNDSWR 180
      | | | | : | | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLVKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | | : | | | | | | | | | | | | | | | |
Db    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSRDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMNRPFLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGKWKIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    293 CCVRQLYIDFRKDLGKWKIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy    361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 5  
A26960  
transforming growth factor beta-1 precursor - green monkey  
C;Species: Cercopithecus aethiops (green monkey, grivet)

C;Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 24-Nov-1999  
 C;Accession: A26960  
 R;Sharples, K.; Plowman, G.D.; Rose, T.M.; Twardzik, D.R.; Purchio, A.F.  
 DNA 6, 239-244, 1987  
 A;Title: Cloning and sequence analysis of simian transforming growth factor-beta  
 cDNA.  
 A;Reference number: A26960; MUID:87246074; PMID:3474130  
 A;Accession: A26960  
 A;Molecule type: mRNA  
 A;Residues: 1-390 <SHA>  
 A;Cross-references: GB:M16658; NID:g176552; PIDN:AAA35369.1; PID:g176553  
 C;Superfamily: inhibin  
 C;Keywords: growth factor  
 F;1-16/Domain: signal sequence #status predicted <SIG>  
 F;17-390/Product: transforming growth factor beta #status predicted <MAT>

Query Match 90.8%; Score 1919; DB 2; Length 390;  
 Best Local Similarity 91.5%; Pred. No. 2.5e-145;  
 Matches 364; Conservative 10; Mismatches 16; Indels 8; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAI R GQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTTPSRPAAGLSTCKTIDMELVKKRIETIR GQILSKLRLA 60

Qy     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    181 YLSNRLLAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHCSCDSKDNTLQVDINGFT 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 TGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRHRRALDTNYCFSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy    361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
  
```

RESULT 6  
 JC4023  
 transforming growth factor beta-1 - dog  
 C;Species: Canis lupus familiaris (dog)  
 C;Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 24-Nov-1999  
 C;Accession: JC4023  
 R;Manning, A.M.; Auchampach, J.A.; Drong, R.F.; Slightom, J.L.  
 Gene 155, 307-308, 1995

A;Title: Cloning of a canine cDNA homologous to the human transforming growth factor-beta 1-encoding gene.  
A;Reference number: JC4023; MUID:95237630; PMID:7721110  
A;Accession: JC4023  
A;Molecule type: mRNA  
A;Residues: 1-390 <MAN>  
A;Cross-references: GB:L34956; NID:g516071; PIDN:AAA51458.1; PID:g516072  
C;Comment: This factor plays a multifunctional role as a regulator of mammalian cell growth and as a modulator of immune responses.  
C;Genetics:  
A;Gene: tgfbeta1  
C;Superfamily: inhibin  
C;Keywords: growth factor; transforming protein  
F;288-390/Product: transforming growth factor beta 1 #status predicted <MAT>

Qy	1	MPSGLRLLPLLLLPLLRLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA	60
		:	
Db	1	MPPSGLRLLPLLLLPLLRLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLS	60
Qy	61	SPPSQGDVPFGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		:   :   :	
Db	61	SPPSQGEVPPVPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVENTNKI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:     :     :	
Db	121	YEKVKKSPHSIYMLFNTSELREAVPEPVLLSRAELRLLRLKLKAQHVELYQKYSNDSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
		:                     : :                     :       :	
Db	181	YLSNRLLAPSDTPEWLSFDVTGVVRQWLSHGGEVEGFRLSAHCSCDSKDNTLQVDINGFS	240
Qy	241	SGRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRRRALDTNYCFSSTDYKDDDDKEKN	300
Db	241	SSRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRQRALDTNYCFSST-----EKN	292
Qy	301	CCVRQLYIDFRKDLGWKIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	360
Db	293	CCVRQLYIDFRKDLGWKIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	352
Qy	361	APCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS	398
Db	353	APCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS	390

A;Title: The murine transforming growth factor-beta precursor.  
A;Reference number: A01396; MUID:86168129; PMID:3007454  
A;Accession: A01396  
A;Molecule type: mRNA  
A;Residues: 1-390 <DER>  
A;Cross-references: GB:M13177; NID:g201952; PIDN:AAA40423.1; PID:g201953  
A;Note: the authors suggest that residues 8-23 could represent the hydrophobic core of an amino-terminal signal peptide  
C;Comment: The mature protein is the carboxyl-terminal segment of a precursor polypeptide; the active molecule is a dimer of identical polypeptide chains linked by an interchain disulfide bond.  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen; transformation  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-278/Domain: propeptide #status predicted <PRO>  
F;244-246/Region: cell attachment (R-G-D) motif  
F;279-390/Product: transforming growth factor beta-1 #status predicted <MAT>  
F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.0%; Score 1840; DB 1; Length 390;  
Best Local Similarity 86.9%; Pred. No. 5e-139;  
Matches 346; Conservative 15; Mismatches 29; Indels 8; Gaps 1;

```

Qy      1 MAPSGLRLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | : | | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YEKTKDISHSIYMFNTSDIREAVPEPPLLSRAELRLQRLKSSVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAI EGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | : | : | | | | | | | | | | | | | | | | | | | | :
Db    181 YLGNRL LTPTDTPWLSFDVTGVVRQWLNQGDGIQGFRFSAHCSDSKDNKLHVEINGIS 240

Qy    241 SGRRGDLATIHGMNRPFLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 PKRRGDLGTIHDNRPFLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy    361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    353 SPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 8  
S10219  
transforming growth factor beta-1 precursor - rat  
N;Alternate names: TGF type 2; TGF-beta





```

Db      241 PKRRGDLGTIHDNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN 292
QY      301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352
QY      361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
        :||||||||||||||||||||||||||||||||||||||
Db      353 SPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 9

A40057

transforming growth factor beta-1 precursor - bovine (fragment)

N;Alternate names: beta-TGF; cartilage-inducing factor-A; EGF-dependent TGF or dEGF; MGF-b; milk growth factor b; TGF-type II

C;Species: Bos primigenius taurus (cattle)

C;Date: 28-Feb-1992 #sequence\_revision 28-Feb-1992 #text\_change 16-Jul-1999

C;Accession: A40057; A42320; A05284; A24322; B61439

R;Van Obberghen-Schilling, E.; Kondaiah, P.; Ludwig, R.L.; Sporn, M.B.; Baker, C.C.

Mol. Endocrinol. 1, 693-698, 1987

A;Title: Complementary deoxyribonucleic acid cloning of bovine transforming growth factor-beta1.

A;Reference number: A40057; MUID:91042552; PMID:3153459

A;Accession: A40057

A;Molecule type: mRNA

A;Residues: 1-315 <VAN>

A;Cross-references: GB:M36271; NID:g163747; PIDN:AAA30778.1; PID:g163748

R;Ogawa, Y.; Schmidt, D.K.; Dasch, J.R.; Chang, R.J.; Glaser, C.B.

J. Biol. Chem. 267, 2325-2328, 1992

A;Title: Purification and characterization of transforming growth factor-beta2.3 and -beta1.2 heterodimers from bovine bone.

A;Reference number: A42320; MUID:92129307; PMID:1733936

A;Accession: A42320

A;Molecule type: protein

A;Residues: 204-209,'X',211-217 <OGA>

R;Roberts, A.B.; Anzano, M.A.; Meyers, C.A.; Wideman, J.; Blacher, R.; Pan, Y.C.E.; Stein, S.; Lehrman, S.R.; Smith, J.M.; Lamb, L.C.; Sporn, M.B.

Biochemistry 22, 5692-5698, 1983

A;Title: Purification and properties of a type beta transforming growth factor from bovine kidney.

A;Reference number: A05284; MUID:84104793; PMID:6607069

A;Accession: A05284

A;Molecule type: protein

A;Residues: 204-218 <ROB>

R;Seyedin, S.M.; Thompson, A.Y.; Bentz, H.; Rosen, D.M.; McPherson, J.M.; Conti, A.; Siegel, N.R.; Galluppi, G.R.; Piez, K.A.

J. Biol. Chem. 261, 5693-5695, 1986

A;Title: Cartilage-inducing factor-A. Apparent identity to transforming growth factor-beta.

A;Reference number: A24322; MUID:86195954; PMID:3754555

A;Accession: A24322

A;Molecule type: protein

A;Residues: 204-233 <SEY>

R;Jin, Y.; Cox, D.A.; Knecht, R.; Raschdorf, F.; Cerletti, N.

J. Protein Chem. 10, 565-575, 1991

A;Title: Separation, purification, and sequence identification of TGF-beta1 and TGF-beta2 from bovine milk.  
A;Reference number: A61439; MUID:92189724; PMID:1799413  
A;Accession: B61439  
A;Molecule type: protein  
A;Residues: 204-209,'X',211-217,'XX',220-232 <JIN>  
C;Comment: This polypeptide is composed of two polypeptide chains cross-linked by disulfide bonds. It has been found in neoplastic and non-neoplastic tissues.  
C;Comment: Type II TGF does not bind to the EGF receptor and lacks intrinsic mitogenic activity, but in soft agar, it reacts synergistically with either type I TGF or EGF, and induces cell proliferation. Cells grown in monolayer do not respond in a similar manner to these growth factors, but morphologically do acquire a transformed phenotype.  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; heterodimer  
F;204-315/Product: transforming growth factor beta-1 #status experimental <MAT>  
F;7,61,101/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.6%; Score 1577; DB 2; Length 315;  
Best Local Similarity 91.3%; Pred. No. 3.5e-118;  
Matches 295; Conservative 9; Mismatches 11; Indels 8; Gaps 1;

```

Qy      76 AVIALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLF 135
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 AILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKIYDKMKSSSHSIYMF 60

Qy     136 NTSELREAVPEPVLLSRaelRLRLKLKVEQHVELYQKYSNDSWRYLsnRLlAPSDSPEW 195
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 NTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWRYLsnRLlAPSDSPEW 120

Qy     196 LSFdVtGVVRQWLtRREAIEGfRLsAHSSSDSKDNTLHVEINGfNSGRRGDLAtIHGMNR 255
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     121 LSFdVtGVVRQWLtRREEIEGfRLsAHCSdSKDNTLQVDINGfSSGRRGDLAtIHGMNR 180

Qy     256 PFLlLMAtPLERaQHLHSSRHRRALDtnYCFsStDYKDDDDKEKnCCVRQLYIDFRKDLG 315
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     181 PFLlLMAtPLERaQHLHSSRHRRALDtnYCFsSt-----EKNCcVRQLYIDFRKDLG 232

Qy     316 WKWiHEPKGYHANfCLGpCPYIwSLDtQYsKVLALYNQHnPGASaAPCCVPQALEPLPIV 375
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     233 WKWiHEPKGYHANfCLGpCPYIwSLDtQYsKVLALYNQHnPGASaAPCCVPQALEPLPIV 292

Qy     376 YYVGRKPKVEQLSNMIVRSCKCS 398
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     293 YYVGRKPKVEQLSNMIVRSCKCS 315

```

# RESULT 10

A41918  
transforming growth factor beta-4 precursor - chicken (fragment)  
N;Alternate names: TGF-beta 4  
C;Species: Gallus gallus (chicken)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
C;Accession: A41918; A34941; S03110  
R;Burt, D.W.; Jakowlew, S.B.  
Mol. Endocrinol. 6, 989-992, 1992

A;Title: Correction: a new interpretation of a chicken transforming growth factor-beta 4 complementary DNA.  
A;Reference number: A41918; MUID:92357039; PMID:1353860  
A;Accession: A41918  
A;Molecule type: mRNA  
A;Residues: 1-373 <BUR>  
A;Cross-references: GB:M31160; GB:X08012; GB:S41706; NID:g1262437; PIDN:AAB05637.1; PID:g1262438  
A;Note: sequence extracted from NCBI backbone (NCBIN:110186, NCBIP:110187)  
A;Note: this report corrects and reinterprets the sequence from reference A34941  
R;Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B.  
Mol. Endocrinol. 2, 1186-1195, 1988  
A;Title: Complementary deoxyribonucleic acid cloning of a messenger ribonucleic acid encoding transforming growth factor beta-4 from chicken embryo chondrocytes.  
A;Reference number: A34941; MUID:89112198; PMID:2464131  
A;Accession: A34941  
A;Molecule type: mRNA  
A;Residues: 'MDPMSIGPKSCGGSPWRPPGTAPWSIGSRRATASSSCSTSSRVRAEVGGRAL',122-209,'D',211-373 <JAK>  
A;Cross-references: EMBL:X08012  
A;Note: this sequence has been corrected in A41918  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor  
F;1/Domain: signal sequence (fragment) #status predicted <SIG>  
F;223-225/Region: cell attachment (R-G-D) motif  
F;260-373/Product: transforming growth factor beta-4 #status predicted <MAT>  
F;54,109,153/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.3%; Score 1063.5; DB 2; Length 373;  
Best Local Similarity 54.9%; Pred. No. 4.1e-77;  
Matches 209; Conservative 51; Mismatches 100; Indels 21; Gaps 8;

```

Qy      30 LSTCKTIDMELVKRKRIEAI RQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVA 89
      ||||: :|:| |:| ||||: ||||| ||||| :|| : || |||: | |||||:: :
Db      2 LSTCQRLDLEAAKKKRIEAVRGQILSKLRLTAPPPASETPPRPLPDDVRALYNSTQELLK 61

Qy      90 GES-VEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPV 148
      : : | |: :|:| |:| : : : : : ||:: :|| | | |
Db      62 QRARLRPPPDGPDEYWAKELRRI PMETTWDGAMEHWQPQSHSIFVFVNSRARRG-GRPT 120

Qy     149 LLSRAELRLLRLKLK-----VEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTGV 203
      || |||||: || | || :||| | | |||| | : : |||||
Db     121 LLHRAELRMLRQKAAADSAGTEQRLELYQGYGNASWRYLHGRSVRATADDEWLSFDVTDA 180

Qy     204 VRQWLTRREAIEGFRLSAHSSSD---SKDNTLHVEINGFNSGRRGDLATI HGMNR--PFL 258
      | |||: | : |:| | : : | || :|||: :| :| |::
Db     181 VHQWLSGSELLGVFKLSVHCPCEMGPGHAEEMRISIEGFEQ-QRGDMQSI AKKHRRVPYV 239

Qy     259 LLMATPLERAQHLHSSRRHRRALDTNYCFS-STDYKDDDDKEKNCCVRQLYIDFRKDLGWK 317
      | || | ||| |||: | |||: || | ||||| ||||| ||||| ||
Db     240 LAMALPAERANELHSARRRRDLDTDYCFGPGTD-----EKNCCVRPLYIDFRKDLQWK 292

Qy     318 WIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYY 377
      ||||| ||||: ||||| ||||: ||||| ||||| ||||| ||||| |||||
Db     293 WIHEPKGYMANFCMGPCPYIWSADTQYTKVLALYNQHNP GASAAPCCVPQTLDPPIIYY 352

```

QY 378 VGRKPKVEQLSNMIVRSCKCS 398  
 ||| : ||||| : || : ||||  
 Db 353 VGRNVRVEQLSNMVVRACKCS 373

RESULT 11

A34939

transforming growth factor beta-3 precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 16-Jul-1999

C;Accession: A34939; S25850; S36125; S36124; I51181

R;Jakowlew, S.B.; Dillard, P.J.; Kondaiah, P.; Sporn, M.B.; Roberts, A.B.

Mol. Endocrinol. 2, 747-755, 1988

A;Title: Complementary deoxyribonucleic acid cloning of a novel transforming growth factor-beta messenger ribonucleic acid from chick embryo chondrocytes.

A;Reference number: A34939; MUID:89096966; PMID:3211158

A;Accession: A34939

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-412 <JAK>

A;Cross-references: GB:M31154; NID:g212758; PIDN:AAA49089.1; PID:g212759

R;Burt, D.W.; Paton, I.R.; Dey, B.R.

J. Mol. Endocrinol. 7, 175-183, 1991

A;Title: Comparative analysis of human and chicken transforming growth factor-beta-2 and -beta-3 promoters.

A;Reference number: S25850; MUID:92134496; PMID:1840616

A;Accession: S25850

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-117 <BUR>

A;Cross-references: EMBL:X58127; NID:g63815; PIDN:CAA41128.1; PID:g63816

A;Accession: S36125

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 119-172 <BU2>

A;Cross-references: EMBL:X60055; NID:g396688; PIDN:CAA42653.1; PID:g396689

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991

A;Accession: S36124

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 173-322, 'ELPT', 327-412 <BU3>

A;Cross-references: EMBL:X60091

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991

R;Jakowlew, S.B.; Lechleider, R.; Geiser, A.G.; Kim, S.J.; Santa-Coloma, T.A.; Cubert, J.; Sporn, M.B.; Roberts, A.B.

Mol. Endocrinol. 6, 1285-1298, 1992

A;Title: Identification and characterization of the chicken transforming growth factor-beta 3 promoter.

A;Reference number: I51181; MUID:93024487; PMID:1406706

A;Accession: I51181

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-117 <JA2>

A;Cross-references: GB:S46000; NID:g257172; PIDN:AAB23575.1; PID:g257173

C;Genetics:

A;Introns: 216/1; 252/1; 309/2; 360/3  
A;Note: list of introns may be incomplete  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-300/Domain: propeptide #status predicted <PRO>  
F;301-412/Product: transforming growth factor beta-3 #status predicted <MAT>  
F;74,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.4%; Score 876; DB 2; Length 412;  
Best Local Similarity 47.1%; Pred. No. 4.3e-62;  
Matches 198; Conservative 56; Mismatches 114; Indels 52; Gaps 14;

```

Qy      15 LLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDVPPGGLP 74
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      9 LVLLSLLSFATVSLALSSCTTLDLEHIKKKRVEAIRGQILSKLRLTSPPE--SVGPAHVP 66

Qy      75 EAVLALYNSTRDRVAGESVEPEPE-----PEADYYAKEVTRVLMVE---SGNQIYDKF 124
      : ||||| : : | : | | | : ||||| : : | : : | : :
Db      67 YQILALYNSTRELL--EEMEEKEESCSQENTESEYYAKEIHKFDMIQGLPEHNELGICP 124

Qy     125 KGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND--- 177
      || : : : || | | | | | : || : : | : | : | |
Db     125 KGVTSNVFR-FNVS---SAEKNSTNLFRAEFRVLRVPNPSSKRSEQRIELFQILRPDEHI 180

Qy     178 -SWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH-----SSSDSKD 229
      |||| | : ||||| |||| | : | : | : | : | : | :
Db     181 AKQRYLSGRNVQTRGSPWLSFDVTDVREWLLHRESNLGLEISIHCPCHTFQPNGDILE 240

Qy     230 N---TLHVEINGFNNSG---RRGDLATI---HGMNRPFLLLMATPLERAQH--LHSSRHRR 278
      | | : : | : | | | : : | : | | | : | : | : |
Db     241 NLHEVLEIKFKGIDSEDDYGRGDLGRLKKQKDLHNPHLILMMLPPHRLESPTLGGQRKKR 300

Qy     279 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 338
      ||||| : | : |||| | ||||| : ||||| : ||||| |||| | |||||
Db     301 ALDTNYCFRNL-----EENCCVRPLYIDFRQDLGWKQVHEPKGYFANFCSGPCPYLR 352

Qy     339 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      | || : | || || | || : |||| | || : |||| | ||||| : | : |||||
Db     353 SADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMIVRSCKCS 412

```

# RESULT 12

B61036

transforming growth factor beta-5 precursor - African clawed frog

C;Species: *Xenopus laevis* (African clawed frog)

C;Date: 31-Dec-1993 #sequence\_revision 03-Feb-1994 #text\_change 16-Jul-1999

C;Accession: A34929; B61036

R;Kondaiah, P.; Sands, M.J.; Smith, J.M.; Fields, A.; Roberts, A.B.; Sporn, M.B.; Melton, D.A.

J. Biol. Chem. 265, 1089-1093, 1990

A;Title: Identification of a novel transforming growth factor-beta (TGF-beta5) mRNA in *Xenopus laevis*.

A;Reference number: A34929; MUID:90110090; PMID:2295601

A;Accession: A34929

A;Molecule type: mRNA

A;Residues: 1-382 <KON>

A;Cross-references: GB:J05180; NID:g214821; PIDN:AAA49968.1; PID:g214822  
R;Roberts, A.B.; Rosa, F.; Roche, N.S.; Coligan, J.E.; Garfield, M.; Rebbert, M.L.; Kondaiah, P.; Danielpour, D.; Kehrl, J.H.; Wahl, S.M.; Dawid, I.B.; Sporn, M.B.  
Growth Factors 2, 135-147, 1990  
A;Title: Isolation and characterization of TGF-beta2 and TGF-beta5 from medium conditioned by Xenopus XTC cells.  
A;Reference number: A61036; MUID:90253806; PMID:2340184  
A;Accession: B61036  
A;Molecule type: protein  
A;Residues: 271-276,'X',278-284,'XX',287-299 <ROB>  
C;Superfamily: inhibin  
C;Keywords: growth factor  
F;271-382/Product: transforming growth factor beta-5 #status experimental <MAT>

Query Match 41.3%; Score 874; DB 2; Length 382;  
Best Local Similarity 46.9%; Pred. No. 5.6e-62;  
Matches 190; Conservative 55; Mismatches 122; Indels 38; Gaps 11;

```

Qy      9 LPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDV 68
      :|  || ||||  |  ||||  :|||  |::| ||||| ||||| :|  |  |
Db      1 MEVLWMLLVLLVLHLSSLAMSLSTCKAVDMEEVRKRRIEAIHQILSKLKLDKTPDV-DS 59

Qy     69 PPGPLPEAVLALYNSTRDRVAGESVEPE-----PEPEADYYAKEVTRVLMVESGNQIYDK 123
      :|  : |||||  : :  :  :  :  : ||||| :|  |  ||  : :|
Db     60 EKMTVPSEAIPLYNSTLEVIREKATREEEHVGHQNIQDYAKQVYR---FESITELED- 115

Qy    124 FKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKL--KVEQHVELYQKYSNDSW-- 179
      |  ||  :||  |  ||  |||| : :  :  : :| :|| :|  :
Db    116 -----HEFKFKFNASHVRENVGMNSLLHHAELRMYKKQTDKNMDQRMELFWKYQENGTT 170

Qy    180 -RYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH---SSSDSKDNTLHVE 235
      ||| : : :|  ||:|||||  | :||  |  |  |  : :||  : :
Db    171 SRYLESKYITPVTDDWMSFDVTKTVNEWLKRAEENEQFGLQPACKCPTPQAKD----ID 226

Qy    236 INGFSNGRRGDLATIHGM--NRPFLLLMTPLERAQHLHSSRHRRALDTNYCFSSTDYKD 293
      ||  ||  ||||| : :  : :| :|  : ||| :|  :  ||| :
Db    227 IEGF-PALRGDLASLSSKENTKPYLMITSMFAERIDTVTSSRKKRGVQGEYCFGNNG--- 282

Qy    294 DDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQ 353
      |||| : |||:||||| ||||| ||: ||| ||||| : ||||| : ||||
Db    283 -----PNCCVKPLYINFRKDLGWKWIHEPKGYEANYCLGNCOPYIWSMDTQYSKVLALYNQ 337

Qy    354 HNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 398
      :||| : ||||  ||||| : ||||  ||||| : ||||  ||
Db    338 NNPGASISPCCVPDVLEPLPIIYVGRKAKVEQLSNMVVRSCNCS 382

```

RESULT 13

S01825

transforming growth factor beta-3 precursor - pig

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 16-Jul-1999

C;Accession: S01825

R;Derynck, R.; Lindquist, P.B.; Lee, A.; Wen, D.; Tamm, J.; Graycar, J.L.; Rhee, L.; Mason, A.J.; Miller, D.A.; Coffey, R.J.; Moses, H.L.; Chen, E.Y.

EMBO J. 7, 3737-3743, 1988

A;Title: A new type of transforming growth factor-beta, TGF-beta3.  
A;Reference number: S01824; MUID:89091120; PMID:3208746  
A;Accession: S01825  
A;Molecule type: mRNA  
A;Residues: 1-409 <DER>  
A;Cross-references: EMBL:X14150; NID:g2127; PIDN:CAA32363.1; PID:g2128  
C;Superfamily: inhibin  
C;Keywords: growth factor  
F;1-25/Domain: signal sequence #status predicted <SIG>  
F;26-297/Domain: propeptide #status predicted <PRO>  
F;298-409/Product: transforming growth factor beta-3 #status predicted <MAT>

Query Match 39.7%; Score 838.5; DB 2; Length 409;  
Best Local Similarity 45.6%; Pred. No. 4.1e-59;  
Matches 190; Conservative 58; Mismatches 122; Indels 47; Gaps 14;

```

QY      15 LLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAI RQILSKLRLASPPSQGDVPPGPLP 74
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      7 LVVLALLNFATVLSMSMSTCTTLD F D H I K R K R V E A I R Q I L S K L R L T S P P D P S M L - - A N I P 64

QY      75 EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMV---ESGNQIYDKFKG 126
      || ||||| : | || : : | : ||||| : : | : | : | : ||
Db      65 TQVLDLYNSTRELLEE V H G E R G D D C T Q E N T E S E Y Y A K E I Y K F D M I Q G L E E H N D L A V C P K G 124

QY      127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND----S 178
      : : || | : : | ||| | : || : : | : | : | : | : |
Db      125 I T S K I F R - F N V S S V E K - - - N E T N L F R A E F R V L R M P N P S S K R S E Q R I E L F Q I L Q P D E H I A K 180

QY      179 WRYLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH-----SSSDSKDN- 230
      || : : | : ||||| || : || || : | : | : | : | : |
Db      181 Q R Y I D G K N L P T R G A A E W L S F D V T D T V R E W L L R R E S N L G L E I S I H C P C H T F Q P N G D I L E N I 240

QY      231 --TLHVEINGFNS---GRRGDLATIHGM--NRPFLLLMATPLERAQH--LHSSRHRRALD 281
      : : : | : | |||| : : | | : || | : | : | : | : |||
Db      241 Q E V M E I K F K G V D S E D D P G R G D L G R L K K K K E H S P H L I L M M I P P D R L D N P G L G A Q R K K R A L D 300

QY      282 TNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD 341
      ||||| : | : ||||| ||||| : ||||| : ||||| : ||||| : ||
Db      301 T N Y C F R N L - - - - - E E N C C V R P L Y I D F R Q D L G W K W V H E P K G Y Y A N F C S G P C P Y L R S A D 352

QY      342 TQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      | : | || ||| || ||||| ||||| : ||||| ||||| : |||||
Db      353 T T H S S V L G L Y N T L N P E A S A S P C C V P Q D L E P L T I L Y Y V G R T A K V E Q L S N M V V K S C K C S 409

```

# RESULT 14

A41397

transforming growth factor beta-3 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 03-Apr-1992 #sequence\_revision 03-Apr-1992 #text\_change 16-Jul-1999

C;Accession: A41397; A61039; A61225

R;Miller, D.A.; Lee, A.; Matsui, Y.; Chen, E.Y.; Moses, H.L.; Derynck, R.

Mol. Endocrinol. 3, 1926-1934, 1989

A;Title: Complementary DNA cloning of the murine transforming growth factor-beta3 (TGFbeta3) precursor and the comparative expression of TGFbeta3 and TGFbeta1 messenger RNA in murine embryos and adult tissues.

A;Reference number: A41397; MUID:90190650; PMID:2628730



A;Accession: A41397  
A;Molecule type: mRNA  
A;Residues: 1-410 <MIL>  
A;Cross-references: GB:M32745; NID:g201949; PIDN:AAA40422.1; PID:g201950  
R;Denhez, F.; Lafyatis, R.; Kondaiah, P.; Roberts, A.B.; Sporn, M.B.  
Growth Factors 3, 139-146, 1990  
A;Title: Cloning by polymerase chain reaction of a new mouse TGF-beta, mTGF-beta3.  
A;Reference number: A61039; MUID:91000714; PMID:2206556  
A;Accession: A61039  
A;Molecule type: mRNA  
A;Residues: 1-410 <DEN>  
R;Watrin, F.; Scotto, L.; Assoian, R.K.; Wolgemuth, D.J.  
Cell Growth Differ. 2, 77-83, 1991  
A;Title: Cell lineage specificity of expression of the murine transforming growth factor beta-3 and transforming growth factor beta-1 genes.  
A;Reference number: A61225; MUID:91299576; PMID:2069871  
A;Accession: A61225  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 285-410 <WAT>  
C;Superfamily: inhibin  
C;Keywords: glycoprotein; growth factor; growth regulation  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-298/Domain: propeptide #status predicted <PRO>  
F;259-261/Region: cell attachment (R-G-D) motif  
F;299-410/Product: transforming growth factor beta-3 #status predicted <MAT>  
F;72,133,140/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 39.4%; Score 832; DB 2; Length 410;  
Best Local Similarity 45.5%; Pred. No. 1.4e-58;  
Matches 190; Conservative 58; Mismatches 122; Indels 48; Gaps 14;

```

Qy      15 LLWLLVLTGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDVPPGGLP 74
      | : | : | : | | | | | : | : | : | : | : | : | : | : | : | : |
Db      7 LVLALLNLATISLSLSTCTTLDFGHIKKRVEAIRGQILSKLRLTSPPEPSVMT--HVP 64

Qy      75 EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMVE---SGNQIYDKFKG 126
      | | | | | | | : : | | : : | : | | | : : | : : | : : | |
Db      65 YQVLALYNSTRELLEEMHGEREEGCTQETSESEYYAKEIHKFDMIQGLAEHNELAVCPKG 124

Qy     127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND----S 178
      : : | | : : | | | | : | : : | : | : | : | : |
Db     125 ITSKVFR-FNVSSVEK---NGTNLFRAEFRVLRVPNPSSKRTEQRIELFQILRPDEHIAK 180

Qy     179 WRYLSNRLLAPSDSPEWLSFDVTVGVVRQWLTRREAIEGFRLSAH-----SSSDSKDN- 230
      | | : : | : | | | | | | : | | | | : | : | : | : |
Db     181 QRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENV 240

Qy     231 --TLHVEINGFNS---GRRGDLATIHGM---NRPFLLMATPLER--AQHLHSSRHRRAL 280
      : : : | : : | | | : : : | : | | | : : | | : | |
Db     241 HEVMEIKFKGVDNEDDHGRGDLGRLKKQKDHHNPHLILMMIPPHRLDSPGQGSQRKKRAL 300

Qy     281 DTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSL 340
      | | | | : : | : | | | | | : | | | : | | | : | | | : |
Db     301 DTNYCFRNL-----EENCCVRPLYIDFRQDLGWKVVHEPKGYANFCSGPCPYLRSA 352

```

Qy 341 DTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398  
 || :| || ||| || |||:||||| |||| | :||||| |||||:|:|||||  
 Db 353 DTHSTVLGLYNTLNPEASAPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKCS 410

# RESULT 15

A55706

transforming growth factor beta-3 precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 16-Jul-1999

C;Accession: A55706; B40699; S36042

R;Wang, J.; Kuliszewski, M.; Yee, W.; Sedlackova, L.; Xu, J.; Tseu, I.; Post, M.  
 J. Biol. Chem. 270, 2722-2728, 1995

A;Title: Cloning and expression of glucocorticoid-induced genes in fetal rat  
 lung fibroblasts. Transforming growth factor-beta-3.

A;Reference number: A55706; MUID:95155340; PMID:7852342

A;Accession: A55706

A;Molecule type: mRNA

A;Residues: 1-410 <WAN>

A;Cross-references: GB:U03491

A;Note: it is uncertain whether Met-1 is the initiator

R;McKinnon, R.D.; Piras, G.; Ida Jr., J.A.; Dubois-Dalcq, M.  
 J. Cell Biol. 121, 1397-1407, 1993

A;Title: A role for TGF-beta in oligodendrocyte differentiation.

A;Reference number: A40699; MUID:93286190; PMID:8509457

A;Accession: B40699

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 157-211 <MCK>

A;Cross-references: EMBL:X71903; NID:g311326; PIDN:CAA50722.1; PID:g311327

C;Superfamily: inhibin

Query Match 39.2%; Score 828; DB 2; Length 410;  
 Best Local Similarity 45.2%; Pred. No. 2.9e-58;  
 Matches 189; Conservative 58; Mismatches 123; Indels 48; Gaps 14;

Qy 15 LLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLASPPSQGDVPPGPLP 74  
 |: |:| : |||| |:| :|:|:||||| ||||| : :|  
 Db 7 LVVLALLNLATVSLSLSTCTTLDGHIKKRVEAIRGQILSKLRLTSPPEPSVMT--HVP 64

Qy 75 EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMVE---SGNQIYDKFKG 126  
 ||||| ||| : : ||| : : ||:||||| : : ||: |||  
 Db 65 YQVLALYNSTRELLEEMHGEREEGCTQETSESEYYAKEIHKFDMIQGLAEHNELAVCPKG 124

Qy 127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND----S 178  
 : : ||| : : |||| |:|: : || :|:|  
 Db 125 ITSKVFR-FNVSSVEK---NGTNLFRAEFRVLRVPNPSSKRTEQRIELFQILRPDEHIAK 180

Qy 179 WRYLSNRLLAPSDSPEWLSFDVTVGVVQWLTRREAIEGFRLSAH-----SSSDSKDN- 230  
 ||: :| : ||||| ||:| |||: | :| :|  
 Db 181 QRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENV 240

Qy 231 --TLHVEINGFNS---GRRGDLATIHGM---NRPFLLLMATPLER--AQHLHSSRRHRAL 280  
 : : :| : : |||| : :| :|:| | :| :|||  
 Db 241 HEVMEIKFKGVNEDDHGRGDLGRLKKQKDHHNPHLILMMIPPHRLDSPGQGGQRKKRAL 300

Qy 281 DTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSL 340

```

      ||||| :      |:||||| |||||:|||||:|||||:||||| |||||: |
Db      301 DTNYCFRNL-----EENCCVRPLYIDFRQDLGWKVVHEPKGYIANFCSGPCPYLRSS 352

Qy      341 DTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      || :| || ||| || |||:||||| ||||| |:||||| ||||| |||||:|:|||||
Db      353 DTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKCS 410

```

Search completed: October 28, 2003, 09:09:48  
 Job time : 14.1517 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 00:08:41 ; Search time 7.97725 Seconds  
 (without alignments)  
 2346.251 Million cell updates/sec

Title: US-10-017-372E-13  
 Perfect score: 2114  
 Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 398

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	2035	96.3	390	1	TGF1_PIG	P07200	sus scrofa
2	1934	91.5	390	1	TGF1_SHEEP	P50414	ovis aries
3	1921	90.9	390	1	TGF1_HUMAN	P01137	homo sapien
4	1919	90.8	390	1	TGF1_CERAE	P09533	cercopithec
5	1898	89.8	390	1	TGF1_CANFA	P54831	canis famil
6	1844	87.2	390	1	TGF1_HORSE	O19011	equus cabal
7	1840	87.0	390	1	TGF1_MOUSE	P04202	mus musculu

8	1840	87.0	390	1	TGF1_RAT	P17246	rattus norv
9	1831	86.6	390	1	TGF1_CAVPO	Q9z1y6	cavia porce
10	1577	74.6	315	1	TGF1_BOVIN	P18341	bos taurus
11	1063.5	50.3	373	1	TGF1_CHICK	P09531	gallus gall
12	874	41.3	382	1	TGF1_XENLA	P16176	xenopus lae
13	857	40.5	412	1	TGF3_CHICK	P16047	gallus gall
14	838.5	39.7	409	1	TGF3_PIG	P15203	sus scrofa
15	832	39.4	410	1	TGF3_MOUSE	P17125	mus musculu
16	828	39.2	412	1	TGF3_RAT	Q07258	rattus norv
17	826	39.1	412	1	TGF3_HUMAN	P10600	homo sapien
18	821.5	38.9	414	1	TGF2_MOUSE	P27090	mus musculu
19	806.5	38.2	414	1	TGF2_HUMAN	P08112	homo sapien
20	806	38.1	413	1	TGF2_XENLA	P17247	xenopus lae
21	805.5	38.1	412	1	TGF2_CHICK	P30371	gallus gall
22	803.5	38.0	435	1	TGF2_PIG	P09858	sus scrofa
23	802.5	38.0	442	1	TGF2_RAT	Q07257	rattus norv
24	468	22.1	112	1	TGF2_BOVIN	P21214	bos taurus
25	322	15.2	375	1	GDF8_MELGA	O42221	meleagris g
26	320	15.1	375	1	GDF8_CHICK	O42220	gallus gall
27	310	14.7	375	1	GDF8_PIG	O18831	sus scrofa
28	309	14.6	375	1	GDF8_PAPHA	O18828	papio hamad
29	307	14.5	376	1	GDF8_RAT	O35312	rattus norv
30	306	14.5	375	1	GDF8_HUMAN	O14793	homo sapien
31	305	14.4	376	1	GDF8_MOUSE	O08689	mus musculu
32	304	14.4	375	1	GDF8_SHEEP	O18830	ovis aries
33	303.5	14.4	405	1	GDFB_MOUSE	Q9z1w4	mus musculu
34	303.5	14.4	407	1	GDFB_HUMAN	O95390	homo sapien
35	295	14.0	375	1	GDF8_BOVIN	O18836	bos taurus
36	289.5	13.7	425	1	IHBA_SHEEP	P43032	ovis aries
37	287	13.6	424	1	IHBA_PIG	P03970	sus scrofa
38	285	13.5	424	1	IHBA_MOUSE	Q04998	mus musculu
39	285	13.5	424	1	IHBA_RAT	P18331	rattus norv
40	284.5	13.5	425	1	IHBA_BOVIN	P07995	bos taurus
41	284	13.4	374	1	GDF8_BRARE	O42222	brachydanio
42	284	13.4	426	1	IHBA_HORSE	P55102	equus cabal
43	283	13.4	426	1	IHBA_HUMAN	P08476	homo sapien
44	282	13.3	345	1	GDFB_RAT	Q9z217	rattus norv
45	280	13.2	424	1	IHBA_CHICK	P27092	gallus gall

# ALIGNMENTS

## RESULT 1

TGF1\_PIG

ID TGF1\_PIG STANDARD; PRT; 390 AA.

AC P07200; P08832;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Transforming growth factor beta 1 precursor (TGF-beta 1).

GN TGFB1.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RX MEDLINE=87174844; PubMed=3470708;  
 RA Derynck R., Rhee L.;  
 RT "Sequence of the porcine transforming growth factor-beta precursor.";  
 RL Nucleic Acids Res. 15:3187-3187(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT VAL-114.  
 RC STRAIN=Miniature swine;  
 RX MEDLINE=89054010; PubMed=2461367;  
 RA Kondaiah P., van Obberghen-Schilling E., Ludwig R.L., Dhar R.,  
 RA Sporn M.B., Roberts A.B.;  
 RT "cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.  
 RT Evidence for alternate splicing and polyadenylation.";  
 RL J. Biol. Chem. 263:18313-18317(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANT VAL-114.  
 RX MEDLINE=88335639; PubMed=3166520;  
 RA Jakowlew S.B., Dillard P.J., Sporn M.B., Roberts A.B.;  
 RT "Nucleotide sequence of chicken transforming growth factor-beta 1  
 RT (TGF-beta 1).";  
 RL Nucleic Acids Res. 16:8730-8730(1988).  
 RN [4]  
 RP SHOWS THAT REF.3 SEQUENCE IS FROM PIG.  
 RA Jakowlew S.B.;  
 RL Unpublished observations (MAR-1996).  
 RN [5]  
 RP SEQUENCE FROM N.A., AND VARIANT VAL-114.  
 RA Wimmers K., Chomdej S., Ponsuksili S., Schellander K.;  
 RT "Polymorphism in the porcine transforming growth factor beta 1  
 RT gene.";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP SEQUENCE OF 279-322.  
 RX MEDLINE=87102890; PubMed=2879635;  
 RA Cheifetz S., Weatherbee J.A., Tsang M.L.S., Anderson J.K., Mole J.E.,  
 RA Lucas R., Massague J.;  
 RT "The transforming growth factor-beta system, a complex pattern of  
 RT cross-reactive ligands and receptors.";  
 RL Cell 48:409-415(1987).  
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -!- CAUTION: REF.3 SEQUENCE WHICH WAS SAID TO ORIGINATE FROM CHICKEN  
 CC WHITE LEGHORN, SEEMS (REF.4) TO ORIGINATE FROM PIG.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; Y00111; CAA68291.1; -.  
DR EMBL; M23703; AAA64616.1; -.  
DR EMBL; X12373; CAA30933.1; -.  
DR EMBL; AF461808; AAL57902.1; -.  
DR PIR; A27512; A27512.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR003911; TGF\_TGFB.  
DR InterPro; IPR001839; TGFB.  
DR InterPro; IPR001111; TGFB\_N.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFB\_propeptide; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRINTS; PR01423; TGFBETA.  
DR ProDom; PD000357; TGFB; 1.  
DR SMART; SM00204; TGFB; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.

KW Growth factor; Mitogen; Glycoprotein; Signal; Polymorphism.

FT	SIGNAL	1	23	POTENTIAL.
FT	PROPEP	24	278	
FT	CHAIN	279	390	TRANSFORMING GROWTH FACTOR BETA 1.
FT	DISULFID	285	294	BY SIMILARITY.
FT	DISULFID	293	356	BY SIMILARITY.
FT	DISULFID	322	387	BY SIMILARITY.
FT	DISULFID	326	389	BY SIMILARITY.
FT	DISULFID	355	355	INTERCHAIN (BY SIMILARITY).
FT	CARBOHYD	82	82	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	176	176	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SITE	244	246	CELL ATTACHMENT SITE (POTENTIAL).
FT	VARIANT	114	114	L -> V.
FT	CONFLICT	6	7	LR -> PG (IN REF. 3).
FT	CONFLICT	180	180	R -> G (IN REF. 3).
FT	CONFLICT	237	237	N -> NA (IN REF. 3).
SQ	SEQUENCE	390 AA;	44294 MW;	A6E2C3659FC384E6 CRC64;

Query Match 96.3%; Score 2035; DB 1; Length 390;

Best Local Similarity 97.0%; Pred. No. 7.2e-153;

Matches 386; Conservative 1; Mismatches 3; Indels 8; Gaps 1;

Qy	1	MAPSGLRLLPLLLPLLWLLVLT	TPGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Db	1	MPPSGLRLLPLLLPLLWLLVLT	TPGRPAAGLSTCKTIDMELVKRKRIE	AIRGQILSKLRLA	60
Qy	61	SPPSQGDVPPGPLPEAVLALYNSTR	DRVAGESVEPEPEPEADYYAKEVTR	VLMLVESGNQI	120
Db	61	SPPSQGDVPPGPLPEAVLALYNSTR	DRVAGESVEPEPEPEADYYAKEVTR	VLMLVESGNQI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAV	PEPVLLSRAELRLLRLKLKVEQHVELY	QKYSNDSWR	180
Db	121	YDKFKGTPHSLYMLFNTSELREAV	PEPVLLSRAELRLLRLKLKVEQHVELY	QKYSNDSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTV	GVVRQWLTRREAI	EGFRLSAHSSSDSKDNTLHVEINGFN	240

```

Db      181 YLSNRLLAPSDSPEWLSFDVTVGVVRQWLTRREAIEGFRLSAHCSCDSKDNTHLVEINGFN 240
QY      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      |||
Db      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN 292
QY      301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      |||
Db      293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352
QY      361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      |||
Db      353 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

## RESULT 2

### TGF1\_SHEEP

```

ID      TGF1_SHEEP      STANDARD;      PRT;      390 AA.
AC      P50414;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 1 precursor (TGF-beta 1).
GN      TGFB1.
OS      Ovis aries (Sheep).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC      Bovidae; Caprinae; Ovis.
OX      NCBI_TaxID=9940;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=95121932; PubMed=7821809;
RA      Woodall C.J., McLaren L.J., Watt N.J.;
RT      "Sequence and chromosomal localisation of the gene encoding ovine
RT      latent transforming growth factor-beta 1.";
RL      Gene 150:371-373(1994).
RN      [2]
RP      SEQUENCE OF 281-390 FROM N.A.
RC      STRAIN=Merino; TISSUE=Skin;
RX      MEDLINE=95268698; PubMed=7749621;
RA      Sutton R., Ward W.G., Raphael K.A., Cam G.R.;
RT      "Growth factor expression in skin during wool follicle development.";
RL      Comp. Biochem. Physiol. 110B:697-705(1995).
CC      -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC      PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC      TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC      HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC      A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC      -!- SUBUNIT: Homodimer; disulfide-linked.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the TGF-beta family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

DR      EMBL; X76916; CAA54242.1; -.
DR      EMBL; L36038; AAA31526.1; -.
DR      PIR; I46463; I46463.
DR      HSSP; P01137; 1KLA.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR003911; TGF_TGFB.
DR      InterPro; IPR001839; TGFB.
DR      InterPro; IPR001111; TGFB_N.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFB_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFBETA.
DR      ProDom; PD000357; TGFB; 1.
DR      SMART; SM00204; TGFB; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
KW      Growth factor; Mitogen; Glycoprotein; Signal.
FT      SIGNAL          1          23          POTENTIAL.
FT      PROPEP          24         278          POTENTIAL.
FT      CHAIN           279         390          TRANSFORMING GROWTH FACTOR BETA 1.
FT      DISULFID        285         294          BY SIMILARITY.
FT      DISULFID        293         356          BY SIMILARITY.
FT      DISULFID        322         387          BY SIMILARITY.
FT      DISULFID        326         389          BY SIMILARITY.
FT      DISULFID        355         355          INTERCHAIN (BY SIMILARITY).
FT      CARBOHYD         82          82          N-LINKED (GLCNAC. . .) (POTENTIAL)
FT      CARBOHYD        136         136          N-LINKED (GLCNAC. . .) (POTENTIAL)
FT      CARBOHYD        176         176          N-LINKED (GLCNAC. . .) (POTENTIAL)
FT      SITE            244         246          CELL ATTACHMENT SITE (POTENTIAL).
SQ      SEQUENCE        390 AA;  44291 MW;  1C247299484D0E57 CRC64;

```

Query Match 91.5%; Score 1934; DB 1; Length 390;  
Best Local Similarity 91.7%; Pred. No. 6.5e-145;  
Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1.

Qy	1	MAPSGRLRLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLRA	60
		:	
Db	1	MPPSGRLRLPLLLPLLWLLMLTPGRP VAGLSTCKTIDMELVKRKGIEAIRGQILSKRLRA	60
Qy	61	SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		:	
Db	61	SPPSQGDVPPGGLPEAILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:     :	
Db	121	YDKMKSSSHSIYMFNTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
Db	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTHREEIEGFRLSAHCSDSKDNTLQVDINGFS	240
Qy	241	SGRRGDLATIHGMRPFLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN	300
Db	241	SGRRGDLATIHGMRPFLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN	292

Db 1 MPPSGLRLLPLLLPLLWLLMLTPGRPVAGLSTCKTIDMELVKRKGIEAIRGOILSKLRLA 60

Ov 61 SPPSOGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVIMVESGNOI 120

Db 61 SPPSOGDVPPGPLPEAILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVIMVEYGNKI 120

Qy 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLKVEOHVELYOKYSNDSWR 180

Db 121 YDKMKSSSHSI YMFFNTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYOKYSNNSWR 180

Qy 181 YLSNRLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240

Db 181 YLSNRLAPSDSPEWLSFDVTGVVRQWLTHREEIEGFRLSAHCSCDSKDNTLQVDINGFS 240

QY 241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300

DB 241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN 292



QY 301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

QY 361 APCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 398  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 353 APCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 3

TGF1\_HUMAN

ID TGF1\_HUMAN STANDARD; PRT; 390 AA.  
 AC P01137; Q9UCG4;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
 GN TGFB1 OR TGFB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87174845; PubMed=3470709;  
 RA Derynck R., Rhee L., Chen E.Y., van Tilburg A.;  
 RT "Intron-exon structure of the human transforming growth factor-beta  
 RT precursor gene.";  
 RL Nucleic Acids Res. 15:3188-3189(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85296301; PubMed=3861940;  
 RA Derynck R., Jarrett J.A., Chen E.Y., Eaton D.H., Bell J.R.,  
 RA Assoian R.K., Roberts A.B., Sporn M.B., Goeddel D.V.;  
 RT "Human transforming growth factor-beta complementary DNA sequence and  
 RT expression in normal and transformed cells.";  
 RL Nature 316:701-705(1985).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Duodenum, and Eye;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 279-390 FROM N.A.  
 RC TISSUE=Carcinoma;  
 RA Urushizaki Y., Niitsu Y., Terui T., Koshida Y., Mahara K., Kohgo Y.,  
 RA Urushizaki I., Takahashi Y., Ito H.;  
 RT "Cloning and expression of the gene for human transforming growth  
 RT factor-beta in Escherichia coli.";  
 RL Tumor Res. 22:41-55(1987).  
 RN [5]  
 RP SEQUENCE OF 279-329.  
 RC TISSUE=Bladder carcinoma;  
 RX MEDLINE=93229900; PubMed=8471846;  
 RA Bourdrel L., Lin C.-H., Lauren S.L., Elmore R.H., Sugarman B.J.,  
 RA Hu S., Westcott K.R.;  
 RT "Recombinant human transforming growth factor-beta 1: expression by  
 RT Chinese hamster ovary cells, isolation, and characterization.";  
 RL Protein Expr. Purif. 4:130-140(1993).  
 RN [6]  
 RP SEQUENCE OF 279-301.  
 RX MEDLINE=85131019; PubMed=2982829;  
 RA Massague J., Like B.;  
 RT "Cellular receptors for type beta transforming growth factor. Ligand  
 RT binding and affinity labeling in human and rodent cell lines.";  
 RL J. Biol. Chem. 260:2636-2645(1985).  
 RN [7]  
 RP STRUCTURE BY NMR OF 279-390.  
 RX MEDLINE=93144319; PubMed=8424942;  
 RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,  
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,  
 RA Torchia D.A.;  
 RT "Transforming growth factor beta 1: NMR signal assignments of the  
 RT recombinant protein expressed and isotopically enriched using Chinese  
 RT hamster ovary cells.";  
 RL Biochemistry 32:1152-1163(1993).  
 RN [8]  
 RP STRUCTURE BY NMR OF 279-390.  
 RX MEDLINE=93144320; PubMed=8424943;  
 RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,  
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,  
 RA Torchia D.A.;  
 RT "Transforming growth factor beta 1: secondary structure as determined  
 RT by heteronuclear magnetic resonance spectroscopy.";  
 RL Biochemistry 32:1164-1171(1993).  
 RN [9]  
 RP STRUCTURE BY NMR OF 279-390.  
 RX MEDLINE=96266150; PubMed=8679613;  
 RA Hinck A.P., Archer S.J., Qian S.W., Roberts A.B., Sporn M.B.,  
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,  
 RA Torchia D.A.;  
 RT "Transforming growth factor beta 1: three-dimensional structure in  
 RT solution and comparison with the X-ray structure of transforming  
 RT growth factor beta 2.";

```

RL   Biochemistry 35:8517-8534(1996).
CC   -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROLS PROLIFERATION,
CC       DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY
CC       CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF THEM HAVE
CC       SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1 REGULATES THE
CC       ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A
CC       POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the TGF-beta family.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X05839; CAA29283.1; -.
DR   EMBL; X05840; CAA29283.1; JOINED.
DR   EMBL; X05843; CAA29283.1; JOINED.
DR   EMBL; X05844; CAA29283.1; JOINED.
DR   EMBL; X05849; CAA29283.1; JOINED.
DR   EMBL; X05850; CAA29283.1; JOINED.
DR   EMBL; X02812; CAA26580.1; ALT_SEQ.
DR   EMBL; BC001180; AAH01180.1; -.
DR   EMBL; BC000125; AAH00125.1; -.
DR   EMBL; BC022242; AAH22242.1; -.
DR   EMBL; M38449; AAA36735.1; -.
DR   PIR; A27513; WFHU2.
DR   PDB; 1KLA; 17-AUG-96.
DR   PDB; 1KLC; 17-AUG-96.
DR   PDB; 1KLD; 17-AUG-96.
DR   Genew; HGNC:11766; TGFB1.
DR   MIM; 190180; -.
DR   GO; GO:0006916; P:anti-apoptosis; TAS.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR003911; TGF_TGFB.
DR   InterPro; IPR001839; TGFB.
DR   InterPro; IPR001111; TGFB_N.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFB_propeptide; 1.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   PRINTS; PR01423; TGFBETA.
DR   ProDom; PD000357; TGFB; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
KW   Growth factor; Mitogen; Glycoprotein; Signal; 3D-structure.
FT   SIGNAL          1      23      POTENTIAL.
FT   PROPEP          24     278
FT   CHAIN           279     390      TRANSFORMING GROWTH FACTOR BETA 1.
FT   DISULFID        285     294
FT   DISULFID        293     356
FT   DISULFID        322     387
FT   DISULFID        326     389
FT   DISULFID        355     355      INTERCHAIN.

```

FT	CARBOHYD	82	82	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	176	176	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SITE	244	246	CELL ATTACHMENT SITE (POTENTIAL).
FT	CONFLICT	10	10	L -> P (IN REF. 2).
FT	CONFLICT	159	159	R -> RR (IN REF. 2).
FT	STRAND	281	281	
FT	TURN	282	287	
FT	STRAND	294	296	
FT	STRAND	300	300	
FT	TURN	302	305	
FT	STRAND	313	313	
FT	STRAND	317	317	
FT	STRAND	321	323	
FT	HELIX	335	346	
FT	TURN	348	349	
FT	STRAND	358	370	
FT	TURN	371	372	
FT	STRAND	373	387	
SQ	SEQUENCE	390 AA;	44341 MW;	75391614250288FE CRC64;

Query Match 90.9%; Score 1921; DB 1; Length 390;  
Best Local Similarity 91.7%; Pred. No. 6.8e-144;  
Matches 365; Conservative 10; Mismatches 15; Indels 8; Gaps 1;

Qy		1	MAPSGRLRLLPLLLPPLLWLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Db		1	MPPSGLRLLLLLLLPLLWLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA	60
Qy		61	SPPSQGDVPPGPPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVL MVESGNQI	120
			:   :   :	
Db		61	SPPSQGEVPPGPPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVL MVETHNEI	120
Qy		121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
			:     :   :	
Db		121	YDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYNNNSWR	180
Qy		181	YLSNRLLAPS DSP E W L S F D V T G V V R Q W L T R R E A I E G F R L S A H S S D S K D N T L H V E I N G F N	240
			:                     :           :	
Db		181	YLSNRLLAPS DSP E W L S F D V T G V V R Q W L S R G G E I E G F R L S A H C S D S R D N T L Q V D I N G F T	240
Qy		241	SGRRGDLATI HG M N R P F L L L M A T P L E R A Q H L H S S R H R R A L D T N Y C F S S T D Y K D D D D K E K N	300
		:		
Db		241	TGRRGDLATI HG M N R P F L L L M A T P L E R A Q H L Q S S R H R R A L D T N Y C F S S T ----- E K N	292
Qy		301	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	360
Db		293	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA	352
Qy		361	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	398
Db		353	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

RESULT 4  
TGF1\_CERAE  
ID TGF1 CERAE STANDARD; PRT; 390 AA.

AC P09533;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Transforming growth factor beta 1 precursor (TGF-beta 1).  
GN TGFB1.  
OS Cercopithecus aethiops (Green monkey) (Grivet).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Cercopithecus.  
OX NCBI\_TaxID=9534;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87246074; PubMed=3474130;  
RA Sharples K., Plowman G.D., Rose T.M., Twardzik D.R., Purchio A.F.;  
RT "Cloning and sequence analysis of simian transforming growth  
RT factor-beta cDNA."  
RL DNA 6:239-244(1987).  
CC -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROL PROLIFERATION,  
CC DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY  
CC CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM HAVE  
CC SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A  
CC POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
CC -!- SUBUNIT: Homodimer; disulfide-linked.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the TGF-beta family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M16658; AAA35369.1; -.  
DR PIR; A26960; A26960.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR003911; TGF\_TGFb.  
DR InterPro; IPR001839; TGFb.  
DR InterPro; IPR001111; TGFb\_N.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFb\_propeptide; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRINTS; PR01423; TGFbeta.  
DR ProDom; PD000357; TGFb; 1.  
DR SMART; SM00204; TGFB; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein; Signal.  
FT SIGNAL 1 16 POTENTIAL.  
FT PROPEP 17 278  
FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
FT DISULFID 285 294 BY SIMILARITY.  
FT DISULFID 293 356 BY SIMILARITY.  
FT DISULFID 322 387 BY SIMILARITY.

Query Match 90.8%; Score 1919; DB 1; Length 390;  
Best Local Similarity 91.5%; Pred. No. 9.8e-144;  
Matches 364; Conservative 10; Mismatches 16; Indels 8; Gaps 1;

Qy	1	MAPSGRLRLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKRLA	60
Db	1	MPPSGRLRLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIETIRGQILSKRLA	60
Qy	61	SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI	120
		:   :	
Db	61	SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEI	120
Qy	121	YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR	180
		:     :	
Db	121	YDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWR	180
Qy	181	YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN	240
		:                           :	
Db	181	YLSNRLLAPSNSPEWLSFDVTGVVRQWLSRGGEIEGFRLSAHSCSDSKDNTLQVDINGFT	240
Qy	241	SGRRGDLATIHGMRNPFLLLMATPLERAQHLHSSRRHRRALDTNYCFSSTDYKDDDDKEKN	300
		:	
Db	241	TGRRGDLATIHGMRNPFLLLMATPLERAQHLQSSRRHRRALDTNYCFSST-----EKN	292
Qy	301	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASA	360
Db	293	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASA	352
Qy	361	APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	398
Db	353	APCCVPOALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	390

## TGF1 CANFA

```

ID      TGF1_CANFA          STANDARD;          PRT;      390 AA.
AC      P54831;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 1 precursor (TGF-beta 1).
GN      TGFBI.
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Jugular vein endothelial;

```

RX MEDLINE=95237630; PubMed=7721110;  
 RA Manning A.M., Auchampach J.A., Drong R.F., Slightom J.L.;  
 RT "Cloning of a canine cDNA homologous to the human transforming growth  
 RT factor-beta 1-encoding gene.";  
 RL Gene 155:307-308(1995).  
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM  
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE  
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES  
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L34956; AAA51458.1; -.  
 DR PIR; JC4023; JC4023.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 278 BY SIMILARITY.  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294 BY SIMILARITY.  
 FT DISULFID 293 356 BY SIMILARITY.  
 FT DISULFID 322 387 BY SIMILARITY.  
 FT DISULFID 326 389 BY SIMILARITY.  
 FT DISULFID 355 355 INTERCHAIN.  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 390 AA; 44185 MW; EB4780E88B7B590E CRC64;

Query Match 89.8%; Score 1898; DB 1; Length 390;  
 Best Local Similarity 90.7%; Pred. No. 4.4e-142;  
 Matches 361; Conservative 12; Mismatches 17; Indels 8; Gaps 1;

Qy 1 MAPSGLRLLPLLLPLLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60





```

CC      -!- SIMILARITY: Belongs to the TGF-beta family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X99438; CAA67801.1; -.
DR      HSSP; P01137; 1KLA.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR003911; TGF_TGFB.
DR      InterPro; IPR001839; TGFB.
DR      InterPro; IPR001111; TGFB_N.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFB_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFBETA.
DR      ProDom; PD000357; TGFB; 1.
DR      SMART; SM00204; TGFB; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
KW      Growth factor; Mitogen; Glycoprotein; Signal.
FT      SIGNAL          1      23      POTENTIAL.
FT      PROPEP          24      278     BY SIMILARITY.
FT      CHAIN           279     390     TRANSFORMING GROWTH FACTOR BETA 1.
FT      DISULFID        285     294     BY SIMILARITY.
FT      DISULFID        293     356     BY SIMILARITY.
FT      DISULFID        322     387     BY SIMILARITY.
FT      DISULFID        326     389     BY SIMILARITY.
FT      DISULFID        355     355     INTERCHAIN (BY SIMILARITY).
FT      CARBOHYD         82      82     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        136     136     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        176     176     N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE       390 AA; 43974 MW; A86D715F44549691 CRC64;

```

```

Query Match          87.2%; Score 1844; DB 1; Length 390;
Best Local Similarity 88.4%; Pred. No. 7.9e-138;
Matches 352; Conservative 12; Mismatches 26; Indels 8; Gaps 1;

```

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RGQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RGQILSKLRLA 60

Qy      61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 SPSPQGEVPPGPLPEAVLALYNSTRAQVAGESAETEPEPEADYYAKEVTRVLMVEKENEI 120

Qy      121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121 YKTVETGSHSIYMFNTSELRAAVDPMLLSRAELRLLRLKLSVEQHVELYQKYSNNSWR 180

Qy      181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181 YLSNRL LTPSDSPEWLSFDVTGVVRQWLSQGGAMEGFRLSAHCSCDSKDNTLRVGINGFS 240

```

Qy	241	SGRRGDLATIHGMRNPFLLLMATPLERAQHLHSSRHRRALDNYCFSSTDYKDDDDKEKN	300
Db	241	SSRRGDLATIDGMNRPFLLLMATPLERAQQHLHSSRHRRALDNYCFSST-----EKN	292
Qy	301	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPGLASA	360
Db	293	CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPGLASA	352
Qy	361	APCCVPQALEPLPIVYYVGRKPKEQLSNMIVRSCKCS	398
Db	353	APCCVPQVLEPLPIVYYVGRKPKEQLSNMIVRSCKCS	390

## RESULT 7

TGF1 MOUSE

ID TGF1 MOUSE STANDARD: PRT: 390 AA.

AC P04202;

DT 20-MAR-1987 (Rel. 04, Created)

DT 20-MAR-1987 (Rel. 04, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Transforming growth factor beta 1 precursor (TGF-beta 1).

GN TGFB1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86168129; PubMed=3007454;

RA Derynck R., Jarrett J.A., Chen E.Y., Goeddel D.V.;

RT "The murine transforming growth factor-beta precursor.";

RL J. Biol. Chem. 261:4377-4379 (1986).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C;

RX MEDLINE=96096545; PubMed=8522200;

RA Guron C., Sudarshan C., Raghov R.;

RT "Molecular organization of the gene encoding murine transforming

```
RT growth factor beta 1.";
```

RL Gene 165:325-326 (1995).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6, and NOD/LT; TISSUE=Spleen;

RA Poirot L., Benoist C., Mathis D.;

RT "Transforming growth factor-beta 1 sequence and expression: no

RT difference between NOD/Lt and C57Bl/6 mouse strains.";

RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL

CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL

CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM

CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE

CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES

CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.

CC -!- SUBUNIT: Homodimer; disulfide-linked.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the TGF-beta family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M13177; AAA40423.1; -.  
 DR EMBL; L42462; AAB00138.1; -.  
 DR EMBL; L42456; AAB00138.1; JOINED.  
 DR EMBL; L42457; AAB00138.1; JOINED.  
 DR EMBL; L42458; AAB00138.1; JOINED.  
 DR EMBL; L42459; AAB00138.1; JOINED.  
 DR EMBL; L42460; AAB00138.1; JOINED.  
 DR EMBL; L42461; AAB00138.1; JOINED.  
 DR EMBL; AJ009862; CAA08900.1; -.  
 DR PIR; A01396; WFMS2.  
 DR HSSP; P01137; 1KLA.  
 DR MGD; MGI:98725; Tgfb1.  
 DR GO; GO:0005578; C:extracellular matrix; IDA.  
 DR GO; GO:0006954; P:inflammatory response; IMP.  
 DR GO; GO:0007515; P:lymph gland development; IMP.  
 DR GO; GO:0008220; P:necrosis; IMP.  
 DR GO; GO:0016202; P:regulation of myogenesis; IDA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 278  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294 BY SIMILARITY.  
 FT DISULFID 293 356 BY SIMILARITY.  
 FT DISULFID 322 387 BY SIMILARITY.  
 FT DISULFID 326 389 BY SIMILARITY.  
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 390 AA; 44310 MW; 4381A51B711D689E CRC64;

Query Match 87.0%; Score 1840; DB 1; Length 390;  
 Best Local Similarity 86.9%; Pred. No. 1.6e-137;  
 Matches 346; Conservative 15; Mismatches 29; Indels 8; Gaps 1;

QY 1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLA 60  
 | |||||

```

Db          1 MPPSGRLRLPLLLPLPWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLA 60
Qy          61 SPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
           |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db          61 SPPSQGEVPPGGLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAI 120
Qy          121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
           |:| | ||:| ||||:||||| ||||| ||||| ||||| ||||| ||||| :|||
Db          121 YEKTKDISHSIYMFNTSDIREAVPEPPLLSRAELRLQRLKSSVEQHVELYQKYSNNSWR 180
Qy          181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
           || |||| |:|:||||| ||||| : : |:| || || || || || || || || :
Db          181 YLGNRLLTPDTPEWLSFDVTGVVRQWLNQGDGIQGFSAHCSDSKDNKLHVEINGIS 240
Qy          241 SGRRGDLATIHGMNRPFLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
           ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          241 PKRRGDLGTIHDNRPFLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN 292
Qy          301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
           ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352
Qy          361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
           :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          353 SPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

# RESULT 8

## TGF1\_RAT

```

ID  TGF1_RAT          STANDARD;          PRT;    390 AA.
AC  P17246;
DT  01-AUG-1990 (Rel. 15, Created)
DT  01-AUG-1990 (Rel. 15, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Transforming growth factor beta 1 precursor (TGF-beta 1).
GN  TGFB1.
OS  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX  NCBI_TaxID=10116;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Sprague-Dawley; TISSUE=Heart;
RX  MEDLINE=90272425; PubMed=2349108;
RA  Qian S.W., Kondaiah P., Roberts A.B., Sporn M.B.;
RT  "cDNA cloning by PCR of rat transforming growth factor beta-1.";
RL  Nucleic Acids Res. 18:3059-3059(1990).
CC  -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC      PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC      TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC      HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC      A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC  -!- SUBUNIT: Homodimer; disulfide-linked.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: Belongs to the TGF-beta family.
CC  -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X52498; CAA36741.1; -.  
 DR PIR; S10219; S10219.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFbBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 278  
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 285 294 BY SIMILARITY.  
 FT DISULFID 293 356 BY SIMILARITY.  
 FT DISULFID 322 387 BY SIMILARITY.  
 FT DISULFID 326 389 BY SIMILARITY.  
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 390 AA; 44329 MW; 5E21108ED50D853C CRC64;

Query Match 87.0%; Score 1840; DB 1; Length 390;  
 Best Local Similarity 86.9%; Pred. No. 1.6e-137;  
 Matches 346; Conservative 14; Mismatches 30; Indels 8; Gaps 1;

QY 1 MAPSGLRLLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAI RGQILSKLRLA 60  
 | |||||  
 Db 1 MPPSGLRLLPLLLPLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAI RGQILSKLRLA 60  
 QY 61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120  
 Db 61 SPSPQGEVPPGPLPEAVLALYNSTRDRVAGESADPEPEPEADYYAKEVTRVLMVDRNNAI 120  
 QY 121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRLKLVKVEQHVELYQKYSNDSWR 180  
 ||| | :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| 180  
 Db 121 YDKTKDITHSIYMFNTSDIREAVPEPPLLSRAELRLQRFKSTVEQHVELYQKYSNNNSWR 180  
 QY 181 YLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240  
 || ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| 240  
 Db 181 YLGNRL LTPTDTPWLSFDVTGVVRQWLNQDGIQGRFRSAHCSDSKDNVLHVEINGIS 240

```

Qy      241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      ||||| ||| ||||||||||||||||||||||||||||||||||||||||| |||
Db      241 PKRRGDLGTIHDMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKN 292

Qy      301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy      361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      : |||||||||||||||||||||||||||||||||||||||
Db      353 SPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

# RESULT 9

## TGF1\_CAVPO

```

ID      TGF1_CAVPO      STANDARD;      PRT;      390 AA.
AC      Q9Z1Y6; Q9QZB3; Q9R148;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 1 precursor (TGF-beta 1).
GN      TGFB1.
OS      Cavia porcellus (Guinea pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX      NCBI_TaxID=10141;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Hartley;
RA      Jeevan A., McMurray D.N., Yoshimura T.;
RT      "Guinea pig transforming growth factor-beta in peritoneal exudates
RT      after BCG vaccination.";
RL      Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN      [2]
RP      SEQUENCE OF 265-382 FROM N.A.
RX      MEDLINE=99144670; PubMed=10025978;
RA      Scarozza A.M., Ramsingh A.I., Wicher V., Wicher K.;
RT      "Spontaneous cytokine gene expression in normal guinea pig blood and
RT      tissues.";
RL      Cytokine 10:851-859(1998).
RN      [3]
RP      SEQUENCE OF 279-371 FROM N.A.
RC      STRAIN=Hartley; TISSUE=Trachea;
RA      Morishima Y., Uchida Y., Nomura A., Ishii Y., Sakamoto T.,
RA      Sekizawa K.;
RT      "Guinea-pig transforming growth factor-beta expression in injured
RT      tracheal epithelium.";
RL      Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC      -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROLS PROLIFERATION,
CC      DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY
CC      CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF THEM HAVE
CC      SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1 REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A
CC      POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC      -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the TGF-beta family.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF191297; AAF02780.1; -.
DR EMBL; AF097509; AAC83807.1; -.
DR EMBL; AF169347; AAD49347.1; -.
DR HSSP; P01137; 1KLA.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR003911; TGF_TGFB.
DR InterPro; IPR001839; TGFB.
DR InterPro; IPR001111; TGFB_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFB_propeptide; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRINTS; PR01423; TGFBETA.
DR ProDom; PD000357; TGFB; 1.
DR SMART; SM00204; TGFB; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Mitogen; Glycoprotein; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 278 POTENTIAL.
FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
FT DISULFID 285 294 BY SIMILARITY.
FT DISULFID 293 356 BY SIMILARITY.
FT DISULFID 322 387 BY SIMILARITY.
FT DISULFID 326 389 BY SIMILARITY.
FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
FT CONFLICT 279 279 G -> P (IN REF. 3).
FT CONFLICT 286 286 F -> S (IN REF. 2).
FT CONFLICT 309 309 K -> E (IN REF. 2).
FT CONFLICT 322 322 C -> R (IN REF. 2).
FT CONFLICT 350 350 A -> G (IN REF. 2).
SQ SEQUENCE 390 AA; 44328 MW; 1539F849BA0C0FF1 CRC64;

```

```

Query Match          86.6%; Score 1831; DB 1; Length 390;
Best Local Similarity 86.7%; Pred. No. 8.4e-137;
Matches 345; Conservative 15; Mismatches 30; Indels 8; Gaps 1;

```

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RGQILSKRLA 60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSRLRLPLLLPLLWLLVLAPGRPASGLSTCKTIDMELVKRKRIEAI RGQILSKRLA 60

Qy     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQI 120
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPPSQGDVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEPDYYAKEVTRVLMVDNSHNI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180

```





CC have been found in bone.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M36271; AAA30778.1; -.  
 DR PIR; A40057; A40057.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein.  
 FT NON\_TER 1 1  
 FT PROPEP <1 203  
 FT CHAIN 204 315 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 210 219 BY SIMILARITY.  
 FT DISULFID 218 281 BY SIMILARITY.  
 FT DISULFID 247 312 BY SIMILARITY.  
 FT DISULFID 251 314 BY SIMILARITY.  
 FT DISULFID 280 280 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 169 171 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 315 AA; 36269 MW; C2717A23D994E00E CRC64;

Query Match 74.6%; Score 1577; DB 1; Length 315;  
 Best Local Similarity 91.3%; Pred. No. 6.3e-117;  
 Matches 295; Conservative 9; Mismatches 11; Indels 8; Gaps 1;

Qy 76 AVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLF 135  
 |:||||||||||||||| | ||||||||||||||||||| |:||| | : ||:| |  
 Db 1 AILALYNSTRDRVAGESAETEPEPEADYYAKEVTRVLMVEYGNKIYDKMKSSSHSIYMFF 60  
 Qy 136 NTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEW 195  
 |||||||||||||||||:|:|||||||||||||||||:|:|||||||||||||||  
 Db 61 NTSELREAVPEPVLLSRADVRLRLKLKVEQHVELYQKYSNNSWRYLSNRLLAPSDSPEW 120  
 Qy 196 LSFDTVGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFNSGRRGDLATIHGMNR 255  
 ||||||||||||||||| ||||||| | ||||||| |:|||:|:|||||||||||  
 Db 121 LSFDTVGVVRQWLTRREEIEGFRLSAHCSCDSKDNTLQVDINGFSSGRRGDLATIHGMNR 180

```

Qy      256 PFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLG 315
      |||||||||||||||||||||||||||||||||||||||
Db      181 PFLLLMATPLERAQHLHSSRHRRALDTNYCFSST-----EKNCCVRQLYIDFRKDLG 232

Qy      316 WKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIV 375
      |||||||||||||||||||||||||||||||||||||||
Db      233 WKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIV 292

Qy      376 YYVGRKPKVEQLSNMIVRSCKCS 398
      ||||||||||||||||||
Db      293 YYVGRKPKVEQLSNMIVRSCKCS 315

```

RESULT 11

TGF1\_CHICK

```

ID      TGF1_CHICK      STANDARD;      PRT;      373 AA.
AC      P09531;
DT      01-MAR-1989 (Rel. 10, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 1 precursor (TGF-beta 1) (TGF-beta 4)
DE      (Fragment).
GN      TGFB1.
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=White leghorn;
RX      MEDLINE=89112198; PubMed=2464131;
RA      Jakowlew S.B., Dillard P.J., Sporn M.B., Roberts A.B.;
RT      "Complementary deoxyribonucleic acid cloning of a messenger
RT      ribonucleic acid encoding transforming growth factor beta 4 from
RT      chicken embryo chondrocytes.";
RL      Mol. Endocrinol. 2:1186-1195(1988).
RN      [2]
RP      REVISIONS.
RX      MEDLINE=92357039; PubMed=1353860;
RA      Burt D.W., Jakowlew S.B.;
RT      "Correction: a new interpretation of a chicken transforming growth
RT      factor-beta 4 complementary DNA.";
RL      Mol. Endocrinol. 6:989-992(1992).
CC      -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC      PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC      TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC      HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC      A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC      -!- SUBUNIT: Homodimer; disulfide-linked.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the TGF-beta family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M31160; AAB05637.1; -.  
DR PIR; A41918; A41918.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR003911; TGF\_TGFb.  
DR InterPro; IPR001839; TGFb.  
DR InterPro; IPR001111; TGFb\_N.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFb\_propeptide; 1.  
DR PRINTS; PR01423; TGFBETA.  
DR ProDom; PD000357; TGFb; 1.  
DR SMART; SM00204; TGFB; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 1 POTENTIAL.  
FT PROPEP 2 259 POTENTIAL.  
FT CHAIN 260 373 TRANSFORMING GROWTH FACTOR BETA 1.  
FT DISULFID 266 277 BY SIMILARITY.  
FT DISULFID 276 339 BY SIMILARITY.  
FT DISULFID 305 370 BY SIMILARITY.  
FT DISULFID 309 372 BY SIMILARITY.  
FT DISULFID 338 338 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT SITE 224 226 CELL ATTACHMENT SITE (POTENTIAL).  
SQ SEQUENCE 373 AA; 42634 MW; 9903F3479C8552E5 CRC64;

Query Match 50.3%; Score 1063.5; DB 1; Length 373;  
Best Local Similarity 54.9%; Pred. No. 2.1e-76;  
Matches 209; Conservative 51; Mismatches 100; Indels 21; Gaps 8;

Qy 30 LSTCKTIDMELVKRKRIEAI RQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVA 89  
| | | | : : | | : | | | | | | | | : | | : | | | | | : :  
Db 2 LSTCQRLDLEAAKKKRIEAVRGQILSKLRLTAPPPASETPPRPLPDDVRALYNSTQELLK 61  
  
Qy 90 GES-VEPEPEPEADYYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPV 148  
: : | | : : | | | | : | : : : : | | : : | | | |  
Db 62 QRARLRPPPDGPDEYWAKELRRIPMETTWGAMEHWQPQSHSIFVFVNSRARRG-GRPT 120  
  
Qy 149 LLSRAELRLLRLKLGK----VEQHVELYQKYSNDSWRYLNRLLAPSDSPEWLSFDVTGV 203  
| | | | | : | | : | | | | | | | | : : | | | | |  
Db 121 LLHRAELRMLRQKAAADSAGTEQRLELYQGYGNASWRYLHGRSVRATADDEWLSFDVTDA 180  
  
Qy 204 VRQWLTRREAIEGFRLSAHSSSD---SKDNTLHVEINGFNSGRRGDLATIHGMNR--PFL 258  
| | | : | : | | | : : : | | : | | : | | : | :  
Db 181 VHQLSGSELLGVFKLSVHCPCEMGPGHAEEMRISIEGFEQ-QRGDMQSIKKHRRVPYV 239  
  
Qy 259 LLMATPLERAQHLHSSRHRRALDTNYCFS-STDYKDDDDKEKNCCVRQLYIDFRKDLGWK 317  
| | | | | | | : | | | | : | | | | | | | | | | | |  
Db 240 LAMALPAERANELHSARRRRDLDTDYCFGPGTD-----EKNCCVRPLYIDFRKDLQWK 292





RESULT 13

TGF3\_CHICK

ID TGF3\_CHICK STANDARD; PRT; 412 AA.  
 AC P16047;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 3 precursor (TGF-beta 3).  
 GN TGFB3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89096966; PubMed=3211158;  
 RA Jakowlew S.B., Dillard P.J., Kondaiah P., Sporn M.B., Roberts A.B.;  
 RT "Complementary deoxyribonucleic acid cloning of a novel transforming  
 RT growth factor-beta messenger ribonucleic acid from chick embryo  
 RT chondrocytes.";  
 RL Mol. Endocrinol. 2:747-755(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=White leghorn;  
 RX MEDLINE=95169270; PubMed=7865129;  
 RA Burt D.W., Dey B.R., Paton I.R., Morrice D.R., Law A.S.;  
 RT "The chicken transforming growth factor-beta 3 gene: genomic  
 RT structure, transcriptional analysis, and chromosomal location.";  
 RL DNA Cell Biol. 14:111-123(1995).  
 RN [3]  
 RP SEQUENCE OF 1-117 FROM N.A.  
 RC STRAIN=White leghorn; TISSUE=Blood;  
 RX MEDLINE=92134496; PubMed=1840616;  
 RA Burt D.W., Dey B.R., Paton I.R.;  
 RT "Comparative analysis of human and chicken transforming growth  
 RT factor-beta 2 and -beta 3 promoters.";  
 RL J. Mol. Endocrinol. 7:175-183(1991).  
 RN [4]  
 RP SEQUENCE OF 1-117 FROM N.A.  
 RX MEDLINE=93024487; PubMed=1406706;  
 RA Jakowlew S.B., Lechleider R., Geiser A.G., Kim S.J.,  
 RA Santa-Coloma T.A., Cubert J., Sporn M.B., Roberts A.B.;  
 RT "Identification and characterization of the chicken transforming  
 RT growth factor-beta 3 promoter.";  
 RL Mol. Endocrinol. 6:1285-1298(1992).  
 CC -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```

CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M31154; AAA49089.1; -.
DR      EMBL; X58127; CAA41128.2; -.
DR      EMBL; X60055; CAA41128.2; JOINED.
DR      EMBL; X60091; CAA41128.2; JOINED.
DR      EMBL; X60090; CAA41128.2; JOINED.
DR      EMBL; S46000; AAB23575.1; -.
DR      PIR; A34939; A34939.
DR      HSSP; P10600; 1TGJ.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR003911; TGF_TGFb.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFBETA.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFB; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
KW      Growth factor; Mitogen; Glycoprotein; Signal.
FT      SIGNAL          1      23      POTENTIAL.
FT      PROPEP          24      300
FT      CHAIN           301      412      TRANSFORMING GROWTH FACTOR BETA 3.
FT      DISULFID        307      316      BY SIMILARITY.
FT      DISULFID        315      378      BY SIMILARITY.
FT      DISULFID        344      409      BY SIMILARITY.
FT      DISULFID        348      411      BY SIMILARITY.
FT      DISULFID        377      377      INTERCHAIN (BY SIMILARITY).
FT      CARBOHYD        74       74      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD       135      135      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD       142      142      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      SITE            261      263      CELL ATTACHMENT SITE (POTENTIAL).
FT      CONFLICT        323      326      ELPT -> DFRQ (IN REF. 1).
SQ      SEQUENCE       412 AA;  47077 MW;  1CAB883170069D55 CRC64;

```

```

Query Match          40.5%;  Score 857;  DB 1;  Length 412;
Best Local Similarity 46.4%;  Pred. No. 4.4e-60;
Matches 195;  Conservative 56;  Mismatches 117;  Indels 52;  Gaps 14;

```

```

QY      15 LLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDVPPGGLP 74
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      9 LVLLSLLSFATVSLALSSCTTLDLEHIKKRVEAIRGQILSKLRLTSPPE--SVGPAHVP 66

QY      75 EAVLALYNSTRDRVAGESVEPEPE-----PEADYYAKEVTRVLMVE---SGNQIYDKF 124
      : | | | | | | | : : | : | | | | : | | | | : : | : : | : :
Db      67 YQILALYNSTRELL--EEMEEEEKEESCSQENTESEYYAKEIHKFDMIQGIPEHNELGICP 124

QY      125 KGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND--- 177
      | | : : : | | | | | | | | | | | | | : | : | : | : | : |
Db      125 KGVTSNVFR-FNVS---SAEKNSTNLFRAEFRVLRVPNPSSKRSEQRIELFQILRPDEHI 180

QY      178 -SWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH-----SSSDSKD 229
      | | | | | : | | | | | | | | : | | | : | : | : | : | :
Db      181 AKQRYLSGRNVQTRGSPEWLSFDVTDTVREWLLHRESNLGLEISIHCPCHTFQPNGDILE 240

```

```

Qy      230 N---TLHVEINGFNSG---RRGDLATI---HGMNRPFLLLMATPLERAQH--LHSSRHRR 278
      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      241 NLHEVLEIKFKGIDSEDDYGRGDLGRLKKQKDLHNPHLILMMLPPHRLESPTLGGQRKKR 300

Qy      279 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 338
      | | | | | | | | :   | : | | | | | | :   | | | | | | : | | | | | | :
Db      301 ALDTNYCFRNL-----EENCCVRPLYIELPTDLGWKQVHEPKGYFANFCSGPCPYLR 352

Qy      339 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      | | | : | | | | | | | | : | | | | | | | | : | | | | | | | | : | | | |
Db      353 SADTTHSTVLGLYNTLNPEASAPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKCS 412

```

RESULT 14

TGF3\_PIG

```

ID      TGF3_PIG          STANDARD;          PRT;    409 AA.
AC      P15203;
DT      01-APR-1990 (Rel. 14, Created)
DT      01-APR-1990 (Rel. 14, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Transforming growth factor beta 3 precursor (TGF-beta 3).
GN      TGFB3.
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX      NCBI_TaxID=9823;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Ovary;
RX      MEDLINE=89091120; PubMed=3208746;
RA      Derynck R., Lindquist B., Lee A., Wen D., Tamm J., Graycar J.L.,
RA      Rhee L., Mason A.J., Miller D.A., Coffey R.J., Moses H.L.,
RA      Chen E.Y.;
RT      "A new type of transforming growth factor-beta, TGF-beta 3.";
RL      EMBO J. 7:3737-3743(1988).
CC      -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.
CC      -!- SUBUNIT: Homodimer; disulfide-linked.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the TGF-beta family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X14150; CAA32363.1; -.
DR      PIR; S01825; S01825.
DR      HSSP; P10600; 1TGJ.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR003911; TGF_TGFB.
DR      InterPro; IPR001839; TGFB.
DR      InterPro; IPR001111; TGFB_N.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFB_propeptide; 1.

```



DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGF BETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT PROPEP 19 297  
 FT CHAIN 298 409 TRANSFORMING GROWTH FACTOR BETA 3.  
 FT DISULFID 304 313 BY SIMILARITY.  
 FT DISULFID 312 375 BY SIMILARITY.  
 FT DISULFID 341 406 BY SIMILARITY.  
 FT DISULFID 345 408 BY SIMILARITY.  
 FT DISULFID 374 374 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 259 261 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 409 AA; 46814 MW; B4900235B5CC955E CRC64;

Query Match 39.7%; Score 838.5; DB 1; Length 409;  
 Best Local Similarity 45.6%; Pred. No. 1.3e-58;  
 Matches 190; Conservative 58; Mismatches 122; Indels 47; Gaps 14;

Qy 15 LLWLLVLTTPGRPAAGLSTCKTIDMELVKRKRIEAI RQILSKLRLASPPSQGDVPPGPLP 74  
 | : | : | : : ||| | : : ||| : ||||| ||||| ||| : : |  
 Db 7 LVVLALLNFATVSLSMSTCTTLD F DHIK RKRVEAIRGQILSKLRLTSPDP SML--ANIP 64  
  
 Qy 75 EAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVLMV---ESGNQIYDKFKG 126  
 || ||||| : | || : : | : ||||| : : | : | : ||  
 Db 65 TQVLDLYNSTRELLEEVHGERGDDCTQENTESEYYAKEIYKFDMIQGLEEHNDLAVCPKG 124  
  
 Qy 127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVEQHVELYQKYSND----S 178  
 : : || | : : | ||| | : || : : || : || : |  
 Db 125 ITSKI FR-FNVSSVEK---NETNLFRAEF RVL RMPNPSSKRSEQRIELFQILQPDEHIAK 180  
  
 Qy 179 WRYLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH-----SSDSKDN- 230  
 || : : | : ||||| ||| : || ||| : | : | : | : |  
 Db 181 QRYIDGKNLPTRGAAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENI 240  
  
 Qy 231 --TLHVEINGFNS---GRRGDLATI HGM--NRPFLLLMATPLERAQH--LHSSRHRRALD 281  
 : : : | : | ||| : : | | : | : | : | : ||||  
 Db 241 QEVM EIKFKGVDEDDPGRGDLGR LKKKKKEHSPHLILMMIPDRLDNPGLGAQRKKRALD 300  
  
 Qy 282 TNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD 341  
 ||||| : : ||||| ||||| : ||||| : ||||| : ||||| : ||  
 Db 301 TNYCFRNL-----EENCCVRPLYIDFRQDLGWKQVHEPKGYANFCSGPCPYLRSAD 352  
  
 Qy 342 TQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398  
 | : | || ||| || ||| : ||||| ||||| : ||||| ||||| : |||||  
 Db 353 TTHSSVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTAKVEQLSNMVVKSCKCS 409

RESULT 15  
 TGF3\_MOUSE  
 ID TGF3\_MOUSE STANDARD; PRT; 410 AA.  
 AC P17125;

DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transforming growth factor beta 3 precursor (TGF-beta 3).  
 GN TGFB3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90190650; PubMed=2628730;  
 RA Miller D.A., Lee A., Matsui Y., Chen E.Y., Moses H.L., Derynck R.;  
 RT "Complementary DNA cloning of the murine transforming growth  
 RT factor-beta 3 (TGF beta 3) precursor and the comparative expression  
 RT of TGF beta 3 and TGF beta 1 messenger RNA in murine embryos and  
 RT adult tissues.";  
 RL Mol. Endocrinol. 3:1926-1934(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91000714; PubMed=2206556;  
 RA Denhez F., Lafyatis R., Kondaiah P., Roberts A.B., Sporn M.B.;  
 RT "Cloning by polymerase chain reaction of a new mouse TGF-beta,  
 RT mTGF-beta 3.";  
 RL Growth Factors 3:139-146(1990).  
 CC -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M32745; AAA40422.1; -.  
 DR PIR; A41397; A41397.  
 DR HSSP; P10600; 1TGJ.  
 DR MGD; MGI:98727; Tgfb3.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR003911; TGF\_TGFB.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 298 POTENTIAL.  
 FT CHAIN 299 410 TRANSFORMING GROWTH FACTOR BETA 3.

Query Match 39.4%; Score 832; DB 1; Length 410;  
Best Local Similarity 45.5%; Pred. No. 4.1e-58;  
Matches 190; Conservative 58; Mismatches 122; Indels 48; Gaps 14;

Search completed: October 28, 2003, 09:08:39  
Job time : 8.97725 secs

OM protein - protein search, using sw model

```
Run on:      October 28, 2003, 07:50:55 ; Search time 31.909 Seconds
              (without alignments)
              3218.683 Million cell updates/sec
```

Title: US-10-017-372E-13  
 Perfect score: 2114  
 Sequence: 1 MAPSGLRLLPLLLPLLWLLV.....GRKPKVEQLSNMIVRSCKCS 398

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SPTREMBL\_23:\*  
 1: sp\_archaea:\*  
 2: sp\_bacteria:\*  
 3: sp\_fungi:\*  
 4: sp\_human:\*  
 5: sp\_invertebrate:\*  
 6: sp\_mammal:\*  
 7: sp\_mhc:\*  
 8: sp\_organelle:\*  
 9: sp\_phage:\*  
 10: sp\_plant:\*  
 11: sp\_rodent:\*  
 12: sp\_virus:\*  
 13: sp\_vertibrate:\*  
 14: sp\_unclassified:\*  
 15: sp\_rvirus:\*  
 16: sp\_bacteriap:\*  
 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	1820	86.1	390	6	Q9TUM8	Q9tum8 equus cabal
2	1741	82.4	368	11	Q8R4D9	Q8r4d9 sigmodon hi
3	848	40.1	379	13	Q8JHF5	Q8jhf5 sparus aura
4	844	39.9	379	13	Q8AXK8	Q8axk8 sparus aura
5	832	39.4	412	11	Q91YU7	Q91yu7 mus musculu
6	825.5	39.0	382	13	Q9PWA9	Q9pwa9 morone chry
7	823.5	39.0	382	13	O93449	O93449 oncorhynchu
8	818.5	38.7	414	11	Q91VP5	Q91vp5 mus musculu
9	788.5	37.3	376	13	Q9PTQ2	Q9ptq2 cyprinus ca
10	757.5	35.8	399	11	Q9ERB7	Q9erb7 mesocricetu
11	724	34.2	362	11	Q99K17	Q99k17 mus musculu

12	696	32.9	130	11	Q08714	Q08714 mesocricetu
13	679	32.1	124	6	Q95N80	Q95n80 canis famil
14	677	32.0	361	13	Q98854	Q98854 cyprinus ca
15	624	29.5	112	6	O02730	O02730 oryctolagus
16	597.5	28.3	255	11	Q921T1	Q921t1 mus musculu
17	589	27.9	127	6	Q9TV08	Q9tv08 canis famil
18	567.5	26.8	224	11	Q8CDZ9	Q8cdz9 mus musculu
19	563	26.6	200	13	Q90YF1	Q90yf1 pleuronecte
20	540	25.5	101	11	Q9R184	Q9r184 meriones un
21	453.5	21.5	179	13	Q90YF2	Q90yf2 pleuronecte
22	413	19.5	88	13	Q90YF5	Q90yf5 pleuronecte
23	402	19.0	88	13	Q90YF7	Q90yf7 oncorhynchu
24	397	18.8	88	13	Q90ZE7	Q90ze7 acipenser b
25	393	18.6	87	13	O42306	O42306 carassius a
26	387	18.3	309	4	Q8WV88	Q8wv88 homo sapien
27	382	18.1	91	6	Q9MYZ1	Q9myz1 capra hircu
28	373	17.6	86	6	Q28241	Q28241 cervus elap
29	358	16.9	81	6	Q9N1S3	Q9n1s3 capreolus c
30	335	15.8	375	13	Q8UWD8	Q8uwd8 columba liv
31	328	15.5	389	13	Q90YY0	Q90yy0 ictalurus p
32	322	15.2	375	13	Q8AVB2	Q8avb2 coturnix co
33	320	15.1	87	13	Q8JHB6	Q8jhb6 scophthalmu
34	320	15.1	375	13	Q8UWD7	Q8uwd7 coturnix ch
35	317	15.0	77	13	Q90YF8	Q90yf8 oncorhynchu
36	315	14.9	375	13	Q98SP0	Q98sp0 gallus gall
37	312	14.8	375	6	Q9GM97	Q9gm97 equus cabal
38	312	14.8	375	13	Q8UWE0	Q8uwe0 anas platyr
39	310	14.7	375	13	Q8UWD9	Q8uwd9 anser anser
40	308	14.6	375	6	Q8HY52	Q8hy52 lepus capen
41	305	14.4	375	6	Q95J86	Q95j86 macaca fasc
42	301.5	14.3	376	13	Q98TB4	Q98tb4 oreochromis
43	301	14.2	62	13	Q90ZJ7	Q90zj7 anguilla an
44	300.5	14.2	376	13	Q90W06	Q90w06 umbrina cir
45	300	14.2	62	13	Q90YF4	Q90yf4 pleuronecte

# ALIGNMENTS

## RESULT 1

Q9TUM8

ID Q9TUM8 PRELIMINARY; PRT; 390 AA.

AC Q9TUM8;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Transforming growth factor beta 1.

GN TGFB1.

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI\_TaxID=9796;

RN [1]

RP SEQUENCE FROM N.A.

RA Nixon A.J., Brower-Toland B.T., Sandell L.J.;

RT "Molecular cloning of equine transforming growth factor beta 1 reveals

RT a divergent nucleotide structure that encodes a novel bioactive

RT peptide among mammalian species.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AF175709; AAD49431.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFbeta.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 SQ SEQUENCE 390 AA; 43860 MW; 220FE40DFCCA6016 CRC64;

Query Match 86.1%; Score 1820; DB 6; Length 390;  
 Best Local Similarity 87.4%; Pred. No. 8.8e-154;  
 Matches 348; Conservative 12; Mismatches 30; Indels 8; Gaps 1;

```

Qy      1 MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIHQILSKLRLA 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIHQILSKLRLA 60

Qy     61 SPSPQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEADYYAKEVTRVLMVESGNI 120
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61 SPSPQGEVPPGPLPEAVLALYNSTRAQVAGESAETEPEPEADYYAKEVTRVLMVEKENEI 120

Qy    121 YDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNDSWR 180
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    121 YKTVETGSHSIYMFNASELRAAVPDPMLLSRAELRLLRLKLSVEQHVELYQKYSNNSWR 180

Qy    181 YLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEINGFN 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    181 YLSNRLLTPSDSPEWLSFDVTGVVRQWLSQGGAMEGLRLSAHCPDCKDNTLRVINGFS 240

Qy    241 SGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKN 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    241 SSRRGDLATIDGMNRPFLLLMATPLERAQQLHSSRHRRALDTNYCSSST-----EKN 292

Qy    301 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    293 CCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP GASA 352

Qy    361 APCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    353 APCCVPQVLEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
  
```

RESULT 2

Q8R4D9

ID Q8R4D9 PRELIMINARY; PRT; 368 AA.  
 AC Q8R4D9;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Transforming growth factor beta-1 protein (Fragment).  
 GN TGFB1.  
 OS Sigmodon hispidus (Hispid cotton rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
 OC Sigmodon.  
 OX NCBI\_TaxID=42415;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Blanco J.C., Pletneva L.M., Prince G.A.;  
 RT "Cotton rat cytokines, chemokines, and interferons.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AF480858; AAL87199.1; -.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFB.  
 DR InterPro; IPR001111; TGFB\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFB\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR ProDom; PD000357; TGFB; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 368 AA; 41905 MW; A5C91207B0468B4A CRC64;

Query Match 82.4%; Score 1741; DB 11; Length 368;  
 Best Local Similarity 86.4%; Pred. No. 9e-147;  
 Matches 325; Conservative 16; Mismatches 27; Indels 8; Gaps 1;

Qy	23	PGRPAAGLSTCKTIDMELVKKRIE	AIRGQILSKLRLASPPSQGDVPPG	PLPEAVLALYN	82
Db	1	PGRPAAGLSTCKTIDMELVKKRIE	AIRGQILSKLRLASPPSQGEVPPG	PLPEAVLALYN	60
Qy	83	STRDRVAGESVEPEPEPEADY	YAKEVTRVLMVESGNQIYDKFKGT	PHSLYMLFNTSELRE	142
		:       :             :			
Db	61	STRDRVAGESADPEPEPEADY	YAKEVTRVLMVDRNNAIYDKTKDI	PHSVYMFNTSDIRE	120
Qy	143	AVPEPVLLSRAELRLRLKLV	EQHVELYQKYSNDSWRYLSNRLL	APSDSPEWLSFDVTG	202
		:     :         :			
Db	121	AVPEPPLLSRAELRLQRFKS	NVEQHVELYEKYSNNSWRYLGN	RLLSPTDSPWLSFDVTS	180
Qy	203	VVRQWLTRREAIEGFRLSA	HSSSDSKDNTLHVEINGFNSGR	RGDLATI HGMNRPFLLLMA	262
		:   :               :			
Db	181	VVRKWLNQGDGIQGFRFSA	HCSDSKDNILHVEINGISPKR	RGDLGTI HDMNRPFLLLMA	240
Qy	263	TPLERAQHLHSSRHRRALD	TNYCFSSTDYKDDDDKEKNCC	VRQLYIDFRKDLGWKWIHEP	322
Db	241	TPLERAQHLHSSRHRRALD	TNYCFSST-----EKNCC	VRQLYIDFRKDLGWKWIHEP	292
Qy	323	KGYHANFCLGPCPYIWSL	TQYSKVLALYNQHNP	GASAAPCCVPQALEPLPIV	382
Db	293	KGYHANFCLGPCPYIWSL	TQYSKVLALYNQHNP	GASASPCCVPQALEPLPIV	352
Qy	383	KVEQLSNMIVRSCKCS	398		

Db 353 KVEQLSNMIVRSCKCS 368

Q8JHF5

Query Match 40.1%; Score 848; DB 13; Length 379;  
Best Local Similarity 46.1%; Pred. No. 4.1e-67;  
Matches 185; Conservative 71; Mismatches 107; Indels 38; Gaps 15;

Qy	12	LLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKKRIEAIRGQILSKLRL-ASPPSQGDVPP	70
		:   : : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	3	LVFLMFMVAYTVGK-VSGMSTCKTLDLEMVKKKRIEAIRSQILSKLRLPTESPOAGD--E	59
Qy	71	GPLPEAVLALYNSTRDRVAGESVEPE-----PEPEADYYAKEVTRVLMVESGNQIYDKFK	125
		:   : :   :   :   : : : :   : : : :   : :   :   : :   :   : :   :	
Db	60	EEIPSSLLSLYNSTKEMLKEQQTEVQTDIFTEQEEEEFYFA---TRVHKFNTTNPV-----	111
Qy	126	GTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLV---EQHVELYQKYSNDSWRYL	182
		: :       :   : :   :   :   :   :   :   :   :   :   :   :   :	
Db	112	RTPQNMSMSFNISEIRRSIGDYRLLTTAELRML-IKAPTILDEQRVELYQGLGT-SPRYL	169
Qy	183	SNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGF--RLSAHSSSDSKDNTLHVEINGFN	240
		: :   : : :   :   :   :   : :   :   : : :   :   :   :   :   :   :   :	



```

Db      170 ASRFITNELRDKWLSFDVTETLQNLWLGNDVQVFQLRLYCDCGRSSDVSTFSFGISGMT 229
Qy      241 SGRRGDLATIHGMNR--PFLLLMATPLERAQHLHSSRHRRALDT-NYCFSSSTDYKDDDDK 297
      :| ||| | : | : |::| |: | : || :|| :|: :| : | : |
Db      230 AG-RGDKAVLDDMTKQPPYILTMSIPKNVSSHL-TSRKKRSTETKDTCTAQTE----- 280
Qy      298 EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP 357
      ||| ||||| ||||| ||||| :| |||:|:| | |||: : :|:|:| ||| |||
Db      281 --TCCVRSLYIDFRKDLGWKWIHKPTRYHANYCMGSCTYIWNAENKYSQILALYKHHNP 338
Qy      358 ASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 398
      || | ||||| ||||| :| ||||| ||||| :| |||||
Db      339 ASAQPCCVQALEPLPILYVGRQHKVEQLSNMIVKSKCS 379

```

#### RESULT 4

Q8AXK8

```

ID   Q8AXK8          PRELIMINARY;          PRT;   379 AA.
AC   Q8AXK8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Transforming growth factor beta 1.
OS   Sparus aurata (Gilthead sea bream).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Sparidae; Sparus.
OX   NCBI_TaxID=8175;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tafalla C., Novoa B., Aranguren R., Figueras A.;
RT   "Molecular cloning and characterization of sea bream (Sparus aurata)
RT   TGF beta 1.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF424703; AAN76665.1; -.
SQ   SEQUENCE   379 AA; 43485 MW; A306EC387F6DBA7C CRC64;

```

```

Query Match          39.9%;   Score 844;   DB 13;   Length 379;
Best Local Similarity 46.1%;   Pred. No. 9.2e-67;
Matches 185;   Conservative 70;   Mismatches 108;   Indels 38;   Gaps 15;

```

```

Qy      12 LLPLLWLLVLTPGRPAAGLSTCKTIDMELVKKRIEAIHQILSKLRL-ASPPSQGDVPP 70
      |:|:::| |: |:| |||||:|:|:| ||||| ||||| | ||
Db      3  LVFLMFMVAYTVGK-VSGMSTCKTLDLEMVKKKRIEAIHQILSKLRLPTESPQAGD--E 59
Qy      71 GPLPEAVLALYNSTRDRVAGESVEPEPE-----PEADYYAKEVTRVLMVESGNQIYDKFK 125
      :| :|:| |||||: : : | : : | :|:| ||| :| :
Db      60 EEIPSSLLSLYNSTKEMLKEQQTEVQTDIFTEXEEEEFYFA---TRVHKFNTTNPV----- 111
Qy      126 GTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKV---EQHVELYQKYSNDSWRYL 182
      || :| | || |:| : : ||: ||||:| :| || |||| | |||
Db      112 RTPQNMSMSFNISEIRRSIGDYRLLTTAELRML-IKAPTILDEQRVELYQGLGT-SPRYL 169
Qy      183 SNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGF--RLSAHSSSDSKDNTLHVEINGFN 240
      ::| : : ||||| :| :| : :| || :| :| :|
Db      170 ASRFITNELRDKWLSFDVTETLQNLWLGNDVQVFQLRLYCDCGRSSDVSTFSFGISGMT 229

```

Qy	241	SGRRGDLATIHGMR--PFLLLMATPLERAQHLHSSRHRRALDT-NYCFSSSTDYKDDDDK	297
		:       :  :   :   :   :   :   :   :	
Db	230	AG-RGDKAVLDDMTKQPPYILTMSIPKNVSSHL-TSRKKRSTETKDTCTAQTE-----	280
Qy	298	EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNP	357
		:         :         :	
Db	281	--TCCVRSLYIDFRKDLGWKWIHKPTRYHANCYCMGSCITYIWNNAENKYSQXLALYKHHNP	338
Qy	358	ASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS	398
		:         :	
Db	339	ASAOPCCVPQALEPLPIVYVGRHKVEOLSNMIVKSCCKCS	379

RESULT 5

Q91YU7

```

ID      Q91YU7      PRELIMINARY;      PRT;      412 AA.
AC      Q91YU7;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Transforming growth factor, beta 3.
GN      TGFB3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Strausberg R.;
RL      Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC      -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR      EMBL; BC014690; AAHL4690.1; -.
DR      MGD; MGI:98727; Tgfb3.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      InterPro; IPR003911; TGF_TGFb.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFbeta.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFB; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
SQ      SEQUENCE      412 AA;  47144 MW;  F3EB65D046DF32AD CRC64;

```

Query Match 39.4%; Score 832; DB 11; Length 412;  
Best Local Similarity 45.5%; Pred. No. 1.2e-65;  
Matches 190; Conservative 58; Mismatches 122; Indels 48; Gaps 14;

```

Qy      15  LLWLLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIRGQILSKLRLASPPSQGDVPPGPLP  74
      | : | : | : | | | | : | : | : | : | | | | | | | | : : |
Db      9  LVVLALLNLATISLSLSTCTTLDGFIKKRVEAIRGQILSKLRLTSPPEPSVMT--HVP  66

Qy     75  EAVLALYNSTR--DRVAGESVE--PEPEPEADYYAKEVTRVLMVE---SGNQIYDKFKG  126
      | | | | | | | | : : | | : : | : | | | | : : | : : | | : |

```

Db 67 YQVLALYNSTRELLEEMHGEREEGCTQETSESEYYAKEIHKFDMIQGLAEHNELAVCPKG 126

Qy 127 TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL----KLKVEQHVELYQKYSND----S 178  
 :: || | : : | || | :||| : || :||| |

Db 127 ITSKVFR-FNVSSVEK---NGTNLFRAEFRVLRVPNPSSKRTEQRIELFQILRPDEHIAK 182

Qy 179 WRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAH-----SSSDSKDN- 230  
 ||: : | : ||||| ||:| |||: | :| | : | :|

Db 183 QRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENV 242

Qy 231 --TLHVEINGFNS---GRRGDLATIHGM---NRPFLLLMATPLER--AQHLHSSRHRRAL 280  
 : : | : : |||| : : | ||| | | : | | :|||

Db 243 HEVMEIKFKGVDNEDDHGRGDLGRLKKQKDHHPHLILMMIPPHRLDSPGQGSQRKKRAL 302

Qy 281 DTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSL 340  
 ||||| : |:|||| |||||:|||||:|||||:|||||:|||||

Db 303 DTNYCFRNL-----EENCCVRPLYIDFRQDLGWKWVHEPKGYANFCSGPCPYLRSA 354

Qy 341 DTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398  
 || :| || ||| || |||:||||| |||| |:||||| |||||:|||:|||||

Db 355 DTHSTVLGLYNTLNPEASAPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKS 412

# RESULT 6

Q9PWA9

ID Q9PWA9 PRELIMINARY; PRT; 382 AA.

AC Q9PWA9;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Transforming growth factor beta precursor.

GN TGF-BETA.

OS Morone chrysops x Morone saxatilis (white bass x striped bass).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;

OC Moronidae; Morone.

OX NCBI\_TaxID=45352;

RN [1]

RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RC TISSUE=KIDNEY;

RX MEDLINE=20394636; PubMed=10938723;

RA Harms C.A., Kennedy-Stoskopf S., Horne W.A., Fuller F.J.,

RA Tompkins W.A.F.;

RT "Cloning and sequencing hybrid striped bass (Morone saxatilis x M.

RT chrysops) transforming growth factor-beta (TGF-beta), and development

RT of a reverse transcription quantitative competitive polymerase chain

RT reaction (RT-qPCR) assay to measure TGF-beta mRNA of teleost fish.";

RL Fish Shellfish Immunol. 10:61-85(2000).

CC -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE

CC RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.

CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).

CC -!- TISSUE SPECIFICITY: HIGHER LEVELS FOUND IN MONONUCLEAR CELLS FROM

CC PERIPHERAL BLOOD THAN IN SPLEEN OR ANTERIOR KIDNEY.

CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

DR EMBL; AF140363; AAD46997.1; -.

DR HSSP; P01137; 1KLA.

```

DR      InterPro; IPRO002400; GF_cysknot.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      InterPro; IPR003911; TGF_TGFb.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFbeta.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFb; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
KW      Growth factor; Mitogen; Glycoprotein; Signal.
FT      SIGNAL          1      ?      POTENTIAL.
FT      PROPEP          ?      270
FT      CHAIN           271     382      TRANSFORMING GROWTH FACTOR BETA.
FT      DISULFID        278     286      BY SIMILARITY.
FT      DISULFID        285     348      BY SIMILARITY.
FT      DISULFID        314     379      BY SIMILARITY.
FT      DISULFID        318     381      BY SIMILARITY.
FT      DISULFID        347     347      INTERCHAIN (BY SIMILARITY).
FT      CARBOHYD        73      73      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD       108     108      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD       113     113      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD       124     124      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD       259     259      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      SITE           235     237      CELL ATTACHMENT SITE (POTENTIAL).
SQ      SEQUENCE       382 AA;  43846 MW;  3124D8C34EA74D72 CRC64;

```

Qy	15	LLWLLVLTPGRPAAGLSTCKTIDMELVKRKRIEAIIRGQILSKLRL---ASPPSQGDVPPG	71
Db	6	LMLVVVYTVGN-VSGMSTCKTLDLEMVKKKRIEAIIRSQILSKLRLPKEPEPDQAGDEEEI	64
Qy	72	PLPEAVLALYNSTRDRVAGESVEPE----PEPEADYYAKEVTRVLMVESGNQIYDKFKG	126
Db	65	PTP--LLSLYNSTKEMLKEQQTEVQTDISTEQEEEEYFAKVLHKKFNMTTRKNN-----	114
Qy	127	TPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLKLKV--EQHVELYQKYSNDSWRYLSN	184
Db	115	TDTTKKMFNISEIRESVGDYRLLTSAELRMLIKKTTIYDEQRVELYSGL-GDSPRYLAS	173
Qy	185	RLLAPSDSPEWLSFDVTVGVVRQWLTRREAIEGF--RLSAHSSSDSKDNTL-HVEINGFNS	241
Db	174	RFITNKWKDKWLSFDVTKTLQDWLKGTDDEQGFQLRLFCECNKVSAGETIFKFGISGIDP	233
Qy	242	GRRGDLATIHGMNR--PFLLLMATPLERAQHLHS---SRHRRALDT-NYCFSSSTDYKDDD	295
Db	234	G-RGDTGPMQLLTQQPPYILTMSIP---QNISSPSTSRKKRSTETKDVCTAQTE-----	283
Qy	296	DKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN	355
Db	284	----TCCVRSLYIDFRKDLGWKWIHKPTGYNANYCMGSCCTYIWAENKYSQILALYKHHN	339
Qy	356	PGASAAPCCVPOALEPLPIVYVYVGRKPKVEQLSNMIVRSCCKS	398

Db 340 PGASAQPCCVPALEPLPILYYVGRQHKVEQLSNMIVKCKCS 382

RESULT 7

O93449

ID O93449 PRELIMINARY; PRT; 382 AA.  
AC O93449; Q91217;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Transforming growth factor beta precursor.  
GN TGF-BETA OR TGF.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LEUKOCYTE;  
RX MEDLINE=99242020; PubMed=10227481;  
RA Daniels G.D., Secombes C.J.;  
RT "Genomic organisation of rainbow trout, Oncorhynchus mykiss TGF-  
RT BETA.";  
RL Dev. Comp. Immunol. 23:139-147(1999).  
RN [2]  
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RC TISSUE=LEUKOCYTE;  
RX MEDLINE=98390168; PubMed=9722928;  
RA Hardie L.J., Laing K.J., Daniels G.D., Grabowski P.S., Cunningham C.,  
RA Secombes C.J.;  
RT "Isolation of the first piscine transforming growth factor beta gene:  
RT analysis reveals tissue specific expression and a potential regulatory  
RT sequence in rainbow trout (Oncorhynchus mykiss).";  
RL Cytokine 10:555-563(1998).  
CC -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE  
CC RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.  
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BLOOD LEUKOCYTES, KIDNEY  
CC MACROPHAGES, BRAIN, GILL AND SPLEEN BUT NOT IN LIVER.  
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
DR EMBL; AJ007836; CAA07707.1; -.  
DR EMBL; X99303; CAA67685.1; -.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR002400; GF\_cysknot.  
DR InterPro; IPR001839; TGFb.  
DR InterPro; IPR001111; TGFb\_N.  
DR InterPro; IPR003911; TGF\_TGFb.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFb\_propeptide; 1.  
DR PRINTS; PR00438; GFCYSKNOT.  
DR PRINTS; PR01423; TGFbeta.  
DR ProDom; PD000357; TGFb; 1.  
DR SMART; SM00204; TGFb; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein; Signal.

FT	SIGNAL	1	20	POTENTIAL.
FT	PROPEP	21	270	
FT	CHAIN	271	382	TRANSFORMING GROWTH FACTOR BETA.
FT	DISULFID	278	286	BY SIMILARITY.
FT	DISULFID	285	348	BY SIMILARITY.
FT	DISULFID	314	379	BY SIMILARITY.
FT	DISULFID	318	381	BY SIMILARITY.
FT	DISULFID	347	347	INTERCHAIN (BY SIMILARITY).
FT	CARBOHYD	76	76	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	116	116	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	125	125	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	237	237	N -> D (IN REF. 2).
FT	CONFLICT	345	345	Q -> H (IN REF. 2).
FT	CONFLICT	371	372	LS -> VP (IN REF. 2).
FT	CONFLICT	377	377	K -> M (IN REF. 2).
SQ	SEQUENCE	382 AA; 44136 MW; 93BD4D3540084B92 CRC64;		

Query Match 39.0%; Score 823.5; DB 13; Length 382;  
 Best Local Similarity 46.9%; Pred. No. 6.3e-65;  
 Matches 180; Conservative 58; Mismatches 107; Indels 39; Gaps 12;

Qy	30	LSTCKTIDMELVKRKRIEAI	RQILSKLRLASPP---	SQGDVPPGPLPEAVLALYNSTRD	86
		:        : :		:    :  : : : :         :	
Db	23	MSTCKSLDLELVKRKRIEAI	RQILSKLRLPKEPEIDQEGDTE--	EVPASLMSIYNSTVE	80
Qy	87	RVAGESVE-----	PEPEPEADYYAKEVTRVLMV	ESGNQIYDKFKGTPHSLYMLFNTSEL	140
		: :     :         :           :         :           :			
Db	81	-LSEEQVHTYIPSTQDAEEEA-	YFAKEVHKFNMKQSENT-----	SKHQI--LFNMSEM	129
Qy	141	REAVPEPVLLSRAELRLL---	RLKLKVEQHVELYQKYSNDS	WRYLSNRL LAPSDSPEWL	196
		:       :                                 :			
Db	130	RSVLGTDRLLSQAELRLLI	KNHGLLDDSEQRLELYRGV-	GDKARYLKSHFVSKewanRWV	188
Qy	197	SFDVTGVVRQWLTRREAIE	GFRLSAHSSSDSKDNTLHVE	INGFN SGRRGDLATI--HGMN	254
		: :     :         :                     :			
Db	189	SFDVTQTLNEWLQGAGEEQ	GFQLKLPDCGKPMEEFRFKI	SGMNK-LRGNTETLAMKMPS	247
Qy	255	RPFLLLMATPLERAQHLH	SSRHRRALDTNYCFSSTDY	KDDDDKEKNCCVRQLYIDFRKDL	314
		:   :       :   :         :   :			
Db	248	KPHILLMSLPVERHSQ	LSSRKKRQTTEEIC-----	SDKSESCVRKLYIDFRKDL	298
Qy	315	GWKWIHEPKGYHANFCL	GPCPYIWSLDTQYSKVLALYN	QHNP GASAAPCCVPQALEPLPI	374
Db	299	GWKWIHEPTGYFANYC	IGPCTYIWNTENKYSQVLALY	KHHNPGASAQPCCVQVLEPLPI	358
Qy	375	VYYVGRKPKVEQLSNMIV	RSCCKS		398
		:         :			
Db	359	IYYVGRQHKVEQLSNMIV	KSCRCs		382

RESULT 8  
 Q91VP5  
 ID Q91VP5 PRELIMINARY; PRT; 414 AA.  
 AC Q91VP5;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Similar to transforming growth factor, beta 2.  
 GN TGFB2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast tumor;  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; BC011170; AAH11170.1; -.  
 DR MGD; MGI:98726; Tgfb2.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 SQ SEQUENCE 414 AA; 47588 MW; DB37A7C38881F286 CRC64;

Query Match 38.7%; Score 818.5; DB 11; Length 414;  
 Best Local Similarity 44.4%; Pred. No. 2e-64;  
 Matches 189; Conservative 61; Mismatches 121; Indels 55; Gaps 16;

Qy 12 LLPLLWLLVLTGPRPAAGLSTCKTIDMELVKKRIEAIHQILSKLRLASPPSQGDVP-P 70  
 :| || | | | ||| |:|: ||||| ||||| ||||| :| || | | |  
 Db 5 VLSTFLLHLVLP--VALSLSTCSTLDMDQFMRKRIEAIHQILSKLKLTSPPPE--DYPEP 60

Qy 71 GPLPEAVIALYNSTRD----RVAGESVEPEPE-PEADYYAKEVTRVLM---VESGNQIYD 122  
 :| |::| |||| : : : | | : : ||||| : : | : | |  
 Db 61 DEVPPEVISIYNSTRDLLQEKASRRAAACERERSDEEYYAKEVYKIDMPSHLPSENAIPP 120

Qy 123 KFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKV-EQHVELYQ-----K 173  
 | | : | : : | : || | : | | | : | | : ||||  
 Db 121 TFY-RPYFRIVRFVDVSTMEKNASN---LVKAEFRVFRQLQNPKEARVAEQRIELYQILKSKD 176

Qy 174 YSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSS----- 225  
 : : : ||: ::: ||||| |::| : : ||:| |  
 Db 177 LTSPTQRYIDSKVVKTRAEGEWLSFDVTDVQEWLHHKDRNLGFKISLHCPCTFVPSNN 236

Qy 226 ---DSKDNTLHVE---INGFNSGRRGDLATIHGMNR-----PFLLLMATPLERAQHLHS 273  
 :| | :| : : || | : : | |||| | | : |  
 Db 237 YIIPNKSEELARFAGIDGTSTYASGDQKTIKSTRKKTSGKTPHLLMLLPYSRLESQQS 296

Qy 274 SRH-RRALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLG 332  
 || : |||| ||| : : |||:| |||||:| ||||| ||||| : |||| |  
 Db 297 SRRKKRALDAAYCFRNV-----QDNCCLRLPLYIDFKRDLGWKWIHEPKGYNANFCAG 348

Qy 333 PCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIV 392  
 |||:| | |||:| ||| || |||:| ||| | |||| | :||:| |||:| |||||

Db 349 ACPYLWSSDTQHTKVLSTYNTINPEASASPCCVSQDLEPLTILYYIGNTPKIEQLSNMIV 408

Qy 393 RSCKCS 398  
:|||||

Db 409 KSCKCS 414

# RESULT 9

## Q9PTQ2

ID Q9PTQ2 PRELIMINARY; PRT; 376 AA.  
AC Q9PTQ2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Transforming growth factor beta precursor.  
OS Cyprinus carpio (Common carp).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Cyprinus.  
OX NCBI\_TaxID=7962;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yin Z., Kuang J.;  
RT "Molecular cloning of carp transforming growth factor beta 1.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE  
CC RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.  
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
DR EMBL; AF136947; AAF22573.1; -.  
DR HSSP; P01137; 1KLA.  
DR InterPro; IPR001839; TGFb.  
DR InterPro; IPR001111; TGFb\_N.  
DR InterPro; IPR003911; TGF\_TGFb.  
DR Pfam; PF00019; TGF-beta; 1.  
DR Pfam; PF00688; TGFb\_propeptide; 1.  
DR PRINTS; PR01423; TGFbeta.  
DR ProDom; PD000357; TGFb; 1.  
DR SMART; SM00204; TGFb; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Growth factor; Mitogen; Glycoprotein; Signal.  
FT SIGNAL 1 22 POTENTIAL.  
FT PROPEP 23 264  
FT CHAIN 265 376 TRANSFORMING GROWTH FACTOR BETA.  
FT DISULFID 272 280 BY SIMILARITY.  
FT DISULFID 308 373 BY SIMILARITY.  
FT DISULFID 312 375 BY SIMILARITY.  
FT DISULFID 341 341 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT SITE 230 232 CELL ATTACHMENT SITE (POTENTIAL).  
SQ SEQUENCE 376 AA; 43329 MW; 7F7FC4DA58B69681 CRC64;

Query Match 37.3%; Score 788.5; DB 13; Length 376;  
Best Local Similarity 44.5%; Pred. No. 8.1e-62;  
Matches 179; Conservative 61; Mismatches 125; Indels 37; Gaps 14;



QY 6 LRLPLLLPLLWLLVLTGRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQ 65  
 :|: ||| | | : ||| :|:|:|:|:|:|:|:|:|:|  
 Db 1 MRVESLLIALQCLLGFV--HYSALSTCSPLDLELIKRRKRIEAIHQILSKLRLSKEPEV 58

QY 66 GDVPPGP-LPEAVLALYNST--RDRVAGESVEPEPEP-EADYYAKEVTR--VLMVESGN 118  
 : :| :::||| : | :|: :| | :||| : : :|  
 Db 59 DEEKESQNIPAE LISVYNSTVELNEEQAAPPEQPKEDPVEEYYAKEVHKFTIKLMEKNP 118

QY 119 QIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRLK--VEQHVELYQKYSN 176  
 ||| : || : : : : :|: ||| | :||| |  
 Db 119 ---DKF-----LWFNITDISQTLGLNRIISQVELRLLITTFPDGSEQRLELYQVIGN 167

QY 177 DSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSAHSSSDSKDNTLHVEI 236  
 | ||| :| : : :||| :| :| | | :|:| :| :| :| :|  
 Db 168 KS-RYLESRFI--PNQRKWLSDVTQTLKDWLQRSEAEQGFQLKMDNCDPQ-KTFQLKI 223

QY 237 NGFNSGRRGDLATIH-GMNRPFLLMATPLERAQHLHSSRRHRALDTNYCFSSTDYKDDD 295  
 | ||| | : | || :|:| :|:| :| :| :| :|  
 Db 224 PGLVL-VRGDTETLAVNMMPRPHILVMSLPLDGNSSSKSRKRQTETDQVC-----T 273

QY 296 DKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 355  
 || ||| | ||||| ||||| ||:|:| | | :|: : :|:| ||| ||  
 Db 274 DKSDGCCVRSLYIDFRKDLGWKWIHEPSGYYANYCTGSCSFVWTSENKYSQVLALYKHHN 333

QY 356 PGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKC 397  
 |||| | ||| | ||| ||||| : ||||| ||||| :|||  
 Db 334 PGASAQPCRVPQVLNPLPIFYVGRQHKVEQLSNMIVKTCKC 375

# RESULT 10

## Q9ERB7

ID Q9ERB7 PRELIMINARY; PRT; 399 AA.  
 AC Q9ERB7;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Transforming growth factor-beta 2 (Fragment).  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 OX NCBI\_TaxID=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ramesh G., Kondaiah P., Seshagiri P.B.;  
 RT "Differential expression and selective localization of transforming  
 RT growth factor-beta isoforms in the hamster uterus during estrous  
 RT cycle.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AY007214; AAG02247.1; -.  
 DR HSSP; P08112; 2TGI.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.

DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR01423; TGFbBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 399 399  
 SQ SEQUENCE 399 AA; 46078 MW; A6FF8E65EAFD5148 CRC64;

Query Match 35.8%; Score 757.5; DB 11; Length 399;  
 Best Local Similarity 43.2%; Pred. No. 5.2e-59;  
 Matches 178; Conservative 58; Mismatches 121; Indels 55; Gaps 16;

QY 18 LLVLTGPRPAAGLSTCKTIDMELVKRKRIEAIHQILSKLRLASPPSQGDVP-PGPLPEA 76  
 || || | ||| |::| | ||||| ||||| | || | || | :|  
 Db 4 LLHLVP--VALSLSTCSTLDMDQFMKRIEAIHQILSKLKLTSPP--DYEPDEVPP 59  
 QY 77 VLALYNSTRD---RVAGESVEPEPE-PEADYYAKEVTRVLM---VESGNQIYDKFKGTP 128  
 |::| |||| | : : | | : : |||| | : | | | | | |  
 Db 60 VISIYNSTRDLLQEASRRAAACERERSDEEYYAKEVYKIDMPSHFPSENAIPTTFY-RP 118  
 QY 129 HSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKV-EQHVELYQ-----KYSNDSW 179  
 : : | : | : : | : || | : | | : | | : || | : : :  
 Db 119 YFRIVRFDVSMMEKNASN--LVKAEFRVFRQLQNPKEARVAEQRIELYQILKSKDLTSPTQ 175  
 QY 180 RYLSNRLLAPSDSPEWLSFDVTGVVQWLTRREAIEGFRLSAH-----SSSDSK 228  
 || : : : : ||||| | : || : : || : | : : : :  
 Db 176 RYIDSKVVKTRAEGEWLSFDVTDVHEWLHHKDRNLGFKISLHCPCTFVPFNNNIIPNK 235  
 QY 229 DNTLHVE---INGFNSGRRGLATIHGMR-----PFLLLMATPLERAQHLHSSRH-RR 278  
 | : | : | : | : | : | : | : | : | : | : | : | :  
 Db 236 SEELEARFAGIDGTSQHSSGHQETIKSTRKKNSGKTPHLLMLLPSYRLESQQSNRRKKR 295  
 QY 279 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 338  
 ||| ||| : : ||| : ||||| : ||||| ||||| ||||| |||||  
 Db 296 ALDAAAYCFRNV-----QDNCCLRPLYIDFRKDLGWKWIHEPKGYNANFCAGACPYLW 347  
 QY 339 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNM 390  
 | ||| : ||| : || | || | : || | || | | : ||| : |||||  
 Db 348 SSDTQHTKVLSTYNTINPEASAPCCVSHDLEPLTILYYIGNTPKIEQLSNM 399

# RESULT 11

Q99K17

ID Q99K17 PRELIMINARY; PRT; 362 AA.  
 AC Q99K17;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Similar to transforming growth factor, beta 3 (Fragment).  
 GN TGFB3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; BC005513; AAH05513.1; -.  
 DR HSSP; P10600; 1TGJ.  
 DR MGD; MGI:98727; Tgfb3.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR InterPro; IPR003911; TGF\_TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR PRINTS; PR01423; TGFbBETA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 362 AA; 41486 MW; 0808E46180FDAE70 CRC64;

Query Match 34.2%; Score 724; DB 11; Length 362;  
 Best Local Similarity 44.3%; Pred. No. 4.4e-56;  
 Matches 166; Conservative 52; Mismatches 109; Indels 48; Gaps 14;

Qy 58 RLASPPSQGDVPPGPLPEAVLALYNSTR---DRVAGESVE--PEPEPEADYYAKEVTRVL 112  
 | : ||| : : | ||||| : : | | : : ||||| : :  
 Db 2 RVGSPPEPSVMT--HVPYQVLALYNSTRELLEEMHGEREEGCTQETSESEYYAKEIHKFD 59  
  
 Qy 113 MVE---SGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL---KLKVE 165  
 | : : | : : | : : | : : | : : | : : | : : | : :  
 Db 60 MIQGLAEHNELAVCPKGITSKVFR-FNVSSVEK--NGTNLFRAEFRVLRVPNPSSKRTE 115  
  
 Qy 166 QHVELYQKYSND----SWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFRLSA 221  
 | : ||| | : : | : : | : ||||| | : || | : | : | :  
 Db 116 QRIELFQILRPDEHIAKQRYIGGKNLPTRGTAEWLSFDVTDTVREWLLRRESNLGLEISI 175  
  
 Qy 222 H-----SSSDSKDN---TLHVEINGFNS---GRRGDLATIHGM---NRPFLLLMATPL 265  
 | : : | : : | : : | : : | : : | : : | : : | : :  
 Db 176 HCPCHTFQPNGDILENVHEVMEIKFKGVDNEDDHGRGDLGRLKKQKDHHNPHLILMMIPP 235  
  
 Qy 266 ER--AQHLHSSRRHRRALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPK 323  
 | : : | : ||||| : : | : |||| | : |||| | : |||| | : ||||  
 Db 236 HRLDSPGQGSQRKKRALDTNYCFRNL-----EENCCVRPLYIDFRQDLGWKWVHEPK 287  
  
 Qy 324 GYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYYVGRKPK 383  
 || : ||| ||||| : | || : | || ||| || ||||| |||| | : ||||| ||  
 Db 288 GYYANFCSGPCPYLRSADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPK 347  
  
 Qy 384 VEQLSNMIVRSCKCS 398  
 ||||| : : |||||  
 Db 348 VEQLSNMVVKSCCKCS 362

RESULT 12  
 Q08714  
 ID Q08714 PRELIMINARY; PRT; 130 AA.  
 AC Q08714; O70331;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).  
 GN TGFB1.  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 OX NCBI\_TaxID=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LVG (SYR);  
 RX MEDLINE=93304479; PubMed=8317544;  
 RA Wong D.T., Donoff R.B., Yang J., Song B.Z., Matossian K., Nagura N.,  
 RA Elovic A., McBride J., Gallagher G., Todd R.;  
 RT "Sequential expression of transforming growth factors alpha and beta 1  
 RT by eosinophils during cutaneous wound healing in the hamster.";  
 RL Am. J. Pathol. 143:130-142(1993).  
 RN [2]  
 RP SEQUENCE OF 26-115 FROM N.A.  
 RC STRAIN=SYRIAN; TISSUE=SPLEEN;  
 RX MEDLINE=98234044; PubMed=9573100;  
 RA Melby P.C., Tryon V.V., Chandrasekar B., Freeman G.L.;  
 RT "Cloning of Syrian hamster (Mesocricetus auratus) cytokine cDNAs and  
 RT analysis of cytokine mRNA expression in experimental visceral  
 RT leishmaniasis.";  
 RL Infect. Immun. 66:2135-2142(1998).  
 CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF  
 CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1  
 CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND  
 CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.  
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
 CC -!- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.  
 DR EMBL; X60296; CAA42838.1; -.  
 DR EMBL; AF046214; AAC40099.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR001839; TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein.  
 FT NON\_TER 1 1  
 FT PROPEP <1 18  
 FT CHAIN 19 130 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 25 34 BY SIMILARITY.  
 FT DISULFID 33 96 BY SIMILARITY.  
 FT DISULFID 66 129 BY SIMILARITY.  
 FT DISULFID 95 95 INTERCHAIN (BY SIMILARITY).  
 FT CONFLICT 93 93 G -> S (IN REF. 2).  
 SQ SEQUENCE 130 AA; 14997 MW; 8B41DD6CF39CCA77 CRC64;

Query Match 32.9%; Score 696; DB 11; Length 130;  
 Best Local Similarity 92.0%; Pred. No. 3.1e-54;

Matches 127; Conservative 0; Mismatches 3; Indels 8; Gaps 1;

```
Qy      261 MATPLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIH 320
      |||
Db      1 MATPLERAQHLQSSRHRRALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIH 52

Qy      321 EPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGR 380
      |||
Db      53 EPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAGPCCVPQALEPLPIVYYVGR 112

Qy      381 KPKVEQLSNMIVRSCKCS 398
      |||
Db      113 KPKVEQLSNMIVRSYKCS 130
```

# RESULT 13

Q95N80

```
ID   Q95N80      PRELIMINARY;      PRT;   124 AA.
AC   Q95N80;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Transforming growth factor beta 1 (Fragment).
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fonfara S., Groene A., Baumgaertner W.;
RT   "Sequence of canine transforming growth factor beta 1 mRNA in DH82-
RT   cells.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AF349538; AAK54072.1; -.
DR   InterPro; IPR001839; TGFb.
DR   Pfam; PF00019; TGF-beta; 1.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
FT   NON_TER      1      1
FT   NON_TER      124    124
SQ   SEQUENCE    124 AA; 14329 MW; 21D185218E5556DB CRC64;
```

Query Match 32.1%; Score 679; DB 6; Length 124;  
Best Local Similarity 93.2%; Pred. No. 9.6e-53;  
Matches 123; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

```
Qy      264 PLERAQHLHSSRHRRALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPK 323
      |||
Db      1 PLERAQHLHSSRQRRALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPK 52

Qy      324 GYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPK 383
      |||
Db      53 GYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPK 112

Qy      384 VEQLSNMIVRSC 395
```

Db                    |||||  
113 VEQLSNMIVRSC 124

RESULT 14

Q98854

ID    Q98854            PRELIMINARY;            PRT;    361 AA.  
AC    Q98854;  
DT    01-FEB-1997 (TrEMBLrel. 02, Created)  
DT    01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT    01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE    Transforming growth factor beta 2 (TGF-beta 2) (Fragment).  
GN    TGFB2.  
OS    Cyprinus carpio (Common carp).  
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC    Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC    Cyprinidae; Cyprinus.  
OX    NCBI\_TaxID=7962;  
RN    [1]  
RP    SEQUENCE FROM N.A.  
RC    TISSUE=HEART;  
RX    MEDLINE=97354301; PubMed=9210595;  
RA    Sumathy K., Desai K.V., Kondaiah P.;  
RT    "Isolation of transforming growth factor-beta2 cDNA from a fish,  
RT    Cyprinus carpio by RT-PCR.";  
RL    Gene 191:103-107(1997).  
CC    -!- FUNCTION: TGF-BETA 2 HAS SUPPRESSIVE EFFECTS ON INTERLEUKIN-2  
CC            DEPENDENT T-CELL GROWTH.  
CC    -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
CC    -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
DR    EMBL; U66874; AAB62983.1; -.  
DR    HSSP; P08112; 2TGI.  
DR    InterPro; IPR001839; TGFb.  
DR    InterPro; IPR001111; TGFb\_N.  
DR    InterPro; IPR003911; TGF\_TGFb.  
DR    Pfam; PF00019; TGF-beta; 1.  
DR    Pfam; PF00688; TGFb\_propeptide; 1.  
DR    PRINTS; PR01423; TGFbeta.  
DR    ProDom; PD000357; TGFb; 1.  
DR    SMART; SM00204; TGFb; 1.  
DR    PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW    Growth factor; Mitogen; Glycoprotein.  
FT    NON\_TER            1            1  
FT    PROPEP            <1          257  
FT    CHAIN            258          361          TRANSFORMING GROWTH FACTOR BETA 2.  
FT    DISULFID        264          273          BY SIMILARITY.  
FT    DISULFID        272          335          BY SIMILARITY.  
FT    DISULFID        334          334          INTERCHAIN (BY SIMILARITY).  
FT    CARBOHYD        30            30          N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT    CARBOHYD        98            98          N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT    CARBOHYD        199          199          N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT    NON\_TER            361          361  
SQ    SEQUENCE        361 AA; 41931 MW; 94D930FA970A3FD3 CRC64;

Query Match                    32.0%;    Score 677;    DB 13;    Length 361;  
Best Local Similarity        41.2%;    Pred. No. 6.7e-52;  
Matches 155; Conservative    54; Mismatches 115; Indels    52; Gaps    14;

Qy 52 QILSKLRLASPPSQGDVPPGP--LPEAVLALYNSTRDRVAGESVEPEPEPE-----ADYY 104  
 ||| ||:| | :| | : :||| | : :| :||  
 Db 1 QILCKLKLSCPP---EIYPEPEEVSRDIIAIYNSTRDLLQEKANERAATCERQRTGEEYY 57

Qy 105 AKEVTRVLM---VESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL- 160  
 |||| :| :| | | : :| :| : :| :||| :||  
 Db 58 AKEVHKIDMQPFYPAENVIPTKHY-NPYFRRLRFDVSSMEKNASN---LVKAELRIFRLQ 113

Qy 161 --KLKV-EQHVELYQ-----KYSNDSWRYLSNRLLAPSDSPEWLSFDVTVGVVRQWLTRRE 212  
 | :| || :||| : : || : : : ||||| | :|| |:  
 Db 114 NPKARVSEQRIELYQILGHKDLTSPTQRYIDSKVVTRTRTEGEWLSFDVTEAVSEWLLHRD 173

Qy 213 AIEGFRLSAHSSS-----DSKDNTLHVEINGFNSG--RRGDLATI----HGMNR 255  
 ||:| | :| | :| :|| |  
 Db 174 RNNGFKISLHCPCTFVPSNNYIIPNKSEELARFAGIDDSFVHGGDLKMFKKRRHSGQS 233

Qy 256 PFLLLMATPLERAQHLHSS-RHRRALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDL 314  
 | ||| | :| | :||| :|| : :||| :||| :|||  
 Db 234 PHLLMLLPSYRLESQHKSHRQKRALDAAFCFRNV-----QDNCCLRSLYIDFKKDL 285

Qy 315 GWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPI 374  
 ||||| :||| | ||:| ||| :| ||| || ||| :||| | ||| |  
 Db 286 GWKWIHEPKGYNANFCAGACPYLWSADTQHSNIGLYNTINPEASASPCCVSDLEPLTI 345

Qy 375 VYYVGRKPKVEQLSNM 390  
 :||:| :||:|  
 Db 346 LYYIGKTPKIEQLSNM 361

# RESULT 15

O02730

ID O02730 PRELIMINARY; PRT; 112 AA.

AC O02730; O97501;

DT 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).

GN TGFB1 OR TGF-BETA-1.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI\_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RA Taylor T.K., James E.R., McGonigle S., Yoho E.R.;

RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.

RN [2]

RP SEQUENCE OF 2-99 FROM N.A.

RA Inoue K., Kawabe Y., Kodama T.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS  
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL  
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF  
 CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1  
 CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND  
 CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.

CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AF000133; AAB53806.1; -.  
 DR EMBL; AB020217; BAA36950.1; -.  
 DR HSSP; P01137; 1KLA.  
 DR InterPro; IPR002400; GF\_cysknot.  
 DR InterPro; IPR001839; TGFb.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR PRINTS; PR00438; GFCYSKNOT.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFB; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Growth factor; Mitogen; Glycoprotein.  
 FT NON\_TER 1 1  
 FT CHAIN 1 112 TRANSFORMING GROWTH FACTOR BETA 1.  
 FT DISULFID 7 16 BY SIMILARITY.  
 FT DISULFID 15 78 BY SIMILARITY.  
 FT DISULFID 44 109 BY SIMILARITY.  
 FT DISULFID 48 111 BY SIMILARITY.  
 FT DISULFID 77 77 INTERCHAIN (BY SIMILARITY).  
 FT CONFLICT 2 3 LD -> FS (IN REF. 2).  
 FT CONFLICT 85 92 PLPIVYYV -> ATAHRTTTL (IN REF. 2).  
 SQ SEQUENCE 112 AA; 12795 MW; 53C5B7D46355A6F3 CRC64;

Query Match 29.5%; Score 624; DB 6; Length 112;  
 Best Local Similarity 93.3%; Pred. No. 6.6e-48;  
 Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 279 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 338  
 |||||  
 Db 1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52  
 Qy 339 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 398  
 |||||  
 Db 53 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112

Search completed: October 28, 2003, 09:12:25  
 Job time : 32.909 secs